Dear Editor,

Please find enclosed our revised manuscript. We have addressed the reviewers' comments as outlined in the responses to the reviewers uploaded earlier.

Overall, we have clarified some points in the text, extended the results section, added new figures and provided supplementary material. All those changes are consistent with referees' suggestions.

The title of the manuscript and some of subsection titles were slightly modified following the suggestions of referee #3.

In the result section, we extended subsections 3.1 and 3.2 and added a new one: *3.3 Validation against ESPI and ONI indices*.
In the new section, we compared two of the most widely used ENSO indices with CESM outputs. In section 3.2, we also added new figures and comments from a new observational data set –GISTEMP- as suggested by referee #4. These new figures highlight the added value of POPEM as suggested by referee 2#. In section 3.1, we clarified the order of the figures and provided more information about the potential applications of POPEM' approach as suggested by referees #1 and #3.

Moreover, the paper was improved with additional references such as those proposed by referee #5.

The revised version of the manuscript also includes other minor changes: updated figure numbers and definition of acronyms.

Attached to this letter you will find a marked-up manuscript version showing all changes to the text. In addition to the text changes, we have also updated and added five new figures and one table. We also included three new figures as supplementary material.

We consider that the manuscript has improved substantially thanks to the work done by the five referees, and hope that it is now acceptable for publication in ESD.

We look forward to your reply.

Andrés Navarro,

Response to referee #1

Referee #1: *The paper is very interesting, novel and merits immediate publication. The approach is that of a 'proof of concept', but the idea behind the research is extremely interesting and worth of attention by the community.*

Reply: Thank you very much. We really appreciate your comments and suggestions.

Referee #1:

However, I believe the authors must touch upon several topics in order to improve the paper. Specifically:

Last part of section 3.1 needs further explanation. Please expand the section and provide more information about potential applications. I think that is an important part of the paper (probably the most important part), and it is a pity that the authors give just such a swift account of the topic.

Reply: Thanks indeed. This is an important point that we didn't explain in full in the first version of the manuscript**.** We added two paragraphs now explaining the potential applications.

The text now reads:

> *Potential applications of POPEM include not only sensitivity analyses of local CO_2 emissions policies, but also the added feature of performing tests for 'what-if' scenarios. One interesting example would be the climate response under the hypothesis that China and India –the most populated countries in the world- reach US CO_2 per capita emissions rates. Another 'what-if' scenario would be the climate response of an increasingly urbanized world. In both cases, POPEM provides a flexible framework for testing the alternative hypotheses.*
>
> *The realism of the ESM will be enhanced with a fully-coupled system. Such a fully-fledged ESM will include bidirectional feedback between POPEM and CESM to evaluate the effects of climate change on population dynamics and emissions.*

Referee #1: *-Given the large number of papers using CESM I think more attention should be devoted to previous work using this model. Please add several references to show how CESM has been used, including merits, shortcomings and the like.*

Reply: We added two paragraphs in section 2.1. They include references from different topics.

The new paragraphs read:

CESM –formerly the Community Climate System Model (CCSM)- was conceived as a coupled atmospheric-oceanic circulation model (Boville and Gent, 1998; Collins et al., 2006; Gent et al., 2011; Hurrell et al., 2013; Williamson, 1983). Since the release of the first version, CESM has evolved into a complex Earth System Model now used in different fields. This includes research into atmospheric (Bacmeister et al., 2014; Liu et al., 2012; Yuan et al., 2013), biogeochemical (Lehner et al., 2015; Nevison et al., 2016; Val Martin et al., 2014), and human-induced processes (Huang and Ullrich, 2016; Levis et al., 2012; Oleson et al., 2011), as well as others. The core code of CESM has also been utilized by various research centers for developing their own models (norESM, Bentsen, 2013; CMCC–CESM–NEMO, Fogli and Iovino, 2014; MIT IGSM-CAM, Monier et al., 2013). CESM has been used in many hundreds of peer-reviewed studies to better understand climate variability and climate change (Hurrell et al., 2013; Kay et al., 2015; Sanderson et al., 2017). Simulations performed with CESM have made a significant contribution to international assessments of climate, including those of the Intergovernmental Panel on Climate Change (IPCC) and the CMIP5/6 project (Coupled Model Intercomparison Project Phase 5/6) (Eyring et al., 2016; IPCC, 2014b; Taylor et al., 2012).

A major advantage of CESM over other ESMs is its availability. Some climate models are developed by scientific groups and access to the source code is limited. The CESM source code is free and available to download from the NCAR website. This approach helps improve the model by setting up a framework for collaborative research and makes the model fully auditable. CESM is a good example of a 'full confidence level' model, after Tapiador et al. (2017), where many 'avatars' of the code are routinely run in several independent research centers, and there is an entire community improving the model and reporting on issues and results. However, the model is not immune to bias. One important shortcoming is the poor representation of precipitation in terms of spatial structure, intensity, duration, and frequency (Dai, 2006; Tapiador et al., 2018; Trenberth et al., 2017, Trenberth et al., 2015). Another major bias is the anomalous warm surface temperature in coastal upwelling regions (Davey et al., 2001; Justin Small, 2015; Richter, 2015).

Referee #1: *-The double ITCZ issue needs referencing. Who did first mention that? Without such reference it seems that that feature is a novel observation from the authors, which I think it is not.*

Reply: Sorry about that. We have added a citation.

The paragraph now reads:

> *The first step in evaluating the new parameterization is to compare the outputs with a control simulation to make sure the new addition does not negatively interact with the dynamical core or spoil the contributions of rest of the parameterizations. Figure 4 shows that this is not case with the POPEM parameterization, which does not negatively affect the outputs of precipitation and temperature. Rather, both variables are now closer to the observed data than they were in the control run, especially in terms of reducing the double ITCZ (Intertropical Convergence Zone), which artificially features in global models (Mechoso et al., 1995; for a recent analysis of double ITCZ in CMIP5 models see Oueslati and Bellon, 2015).*

Referee #1: *-The following sentence is confusing to me. The improvements of POPEM for the El Niño-4 area show that detailed, dynamical modeling of GHG emissions is important for more precisely quantifying precipitation in dry areas, which validates the main hypothesis of the paper. Please explain what do you mean by that.*

Reply: What we meant was that precipitation in dry areas is extremely important, since human activities and biota are highly dependent of it. Improving the representation of precipitation in models is thus crucial. The main hypothesis of the paper, namely that point-wise emissions can improve the modeling, is validated for the El Niño-4 area where we show that our model improves the representation of precipitation in the left tail of the distribution (cf. Figure 8). We have reworded the paragraph:

> *"The results for the El Niño-4 area show that detailed, grid-point emissions of GHG improves the quantification of precipitation in dry areas, in agreement with our hypothesis about the benefits of locally-distributed versus global mean forcings."*

Referee #1: *- Please, check the references to the figures in the text (Figs. 4 and 5)*

Reply: Amended now, thanks.

Referee #1: *- Please, check the place where you put the definition of some acronyms (e. g. ITCZ -you use it on page 7 and is defined in page 10-, SST -similar-).*

Reply: Sorry about that. Amended now.

References

Mechoso, C. R., Robertson, A. W., Barth, N., Davey, M. K., Delecluse, P., Gent, P. R., Ineson, S., Kirtman, B., Latif, M., Treut, H. Le, Nagai, T., Neelin, J. D., Philander, S. G. H., Polcher, J., Schopf, P. S., Stockdale, T., Suarez, M. J., Terray, L., Thual, O. and Tribbia, J. J.: The Seasonal Cycle over the Tropical Pacific in Coupled Ocean–Atmosphere General Circulation Models, Mon. Weather Rev., 123(9), 2825–2838, doi:10.1175/1520-0493(1995)123<2825:TSCOTT>2.0.CO;2, 1995.

Meehl, G. A. and Arblaster, J. M.: The Asian-Australian monsoon and El Nino-Southern Oscillation in the NCAR climate system model, J. Clim., 11(6), 1356–1385, doi:10.1175/1520-0442(1998)011<1356:TAAMAE>2.0.CO;2, 1998.

Oueslati, B. and Bellon, G.: The double ITCZ bias in CMIP5 models: interaction between SST, large-scale circulation and precipitation, Clim. Dyn., 44(3–4), 585–607, doi:10.1007/s00382-015-2468-6, 2015.

Terray, L.: Sensitivity of Climate Drift to Atmospheric Physical Parameterizations in a Coupled Ocean–Atmosphere General Circulation Model, J. Clim., 11(7), 1633–1658, doi:10.1175/1520-0442(1998)011<1633:SOCDTA>2.0.CO;2, 1998.

Response to referee #2

Referee #2:

General Comments:

This paper highlights the importance of grid point scale modeling of anthropogenic pollutants, especially CO2, and the integration of such modeling through a new module called "Population Parameterization for Earth Models" (POPEM). The module is integrated into a highly distributed climate model like Community Earth System Model (CESM). The authors present clearly and adequately the added value of their contribution (POPEM) to the model simulations and underline its impact to the climate predictions of both precipitation and temperature.

Reply: Thanks for your positive feedback.

Referee #2:

Minor comments:

Figures 6B and 8B do not illustrate clearly any differences between GPCP – CONTROL, GPCP – POPEM and CRU – CONTROL, CRU – POPEM respectively. Maybe the authors should consider an alternative way to show the differences.

Reply: Thanks for your suggestion. We added new more detailed figures (Figures 9, 11, 12, 13 and 14) to highlight the added value of our approach.

Response to referee #3. Svetla Hristova-Veleva

Referee #3:

Overview and recommendations:

The paper addresses important questions regarding improving the performance of Earth System Models (ESM) – the important tools to study and understand the complexity of the Earth's climate. Improving these models is a major goal of the science community as they can be a very valuable tool in studying the response of the Earth's system to anthropogenic forcing, providing guidance to policy makers.

Reply: Thanks.

Referee #3:

In particular, the paper investigates the impact of a new parameterization of CO2 emissions that the authors have recently developed, called the POPEM module (POpulation Parameterization for Earth Models). POPEM presents an important advancement in the way CO2 emissions are modeled, as it accounts dynamically for the changing emissions. Like previous research, POPEM uses population data as proxies for emission. What is unique to this new parameterization, though, is that it models the evolution of the population while previous research has relied on historical data, hence not being dynamical, preventing them from making reliable predictions for the future emissions and the response of the climate system. Using this new parameterization (POPEM) presents an important advancement and this makes the described research very valuable.

Reply: Thanks.

However, before going forward one have to evaluate the performance and assess the impact of the new parameterization. Indeed, this is the goal of this paper.

The paper begins by describing what is unique about POPEM.

It then validates the stand-alone performance of POPEM by comparing its predication over a past 63 (and 70) -year period to existing data. The comparison is done globally but also by several regions. This validation is done in two ways: by comparing forecasted to observed population growth rates; and by comparing the forecasted to observed emission rates. The results show that despite the difficulty of predicting non-linear trends in the growth of population and emissions, *POPEM preforms quite well. These comparisons give credibility to the POPEM forecasts, hence to its use in forecasting future scenarios.*

Next, the paper uses a coupled ESM, the Community ESM (CESM) to evaluate the impact of POPEM. The evaluation focusses on the impact of POPEM on two very important, and difficult to predict, parameters of the Earth's system - the precipitation and the sea surface temperature (SST). The evaluation is done in two ways:

- *by comparing the results from a control run (using global CO2 concentration parameters that I believe are homogeneous – this needs clarification) to those from POPEM. This choice of model setups highlights the value of POPEM as it predicts the population (and the emissions) in every grid point, showing the impact and the importance of the spatial variability.*
- *By comparing both control and POPEM forecasts to actual observations (over a 20-year period for precipitation and 50-year period for SST).*

The paper finds that:

- *The global predictions for both parameters compare to the observations in a very similar way for the CONTROL and the POPEM simulations. Hence, the more realistic POPEM parameterization "does no harm". This is an important test and conclusion because it is occasionally the case that including more realistic parameterizations might degrade the performance of the forecasts for certain parameters. This is because often the models are "tuned" to predicting some of the parameters, giving the right answer for the wrong reason, and impacting negatively the forecasting of the non-tuned parameters when the more realistic parameterizations are employed.*
- *More importantly, the paper finds that using POPEM results in regional differences between its forecasts and that of the control run. Comparison to observations seems to suggest the POPEM produces better regional distribution of the precipitation. This is a very important conclusion, in my view. It does not seem to be well highlighted in the paper summary.*

Overall, the paper addresses a very important topic. The approach is sound and uses a very good modeling framework. There is a very extensive set of references. The paper is presented in a fluent and precise language. However, there are several places where the paper could be improved, as detailed below.

Because of all that, I propose the paper be accepted with minor revisions.

Reply: Thanks for highlighting the main findings of the manuscript and for your detailed revision of the paper. Also, thanks for your suggestions and comments. We consider that they improve the global quality of the paper.

Referee #3:

Title: The current title is: "Improving the representation of anthropogenic CO_2 emissions in climate models: a new parameterization for the Community Earth System Model (CESM)"

I would suggest a modification to read "Improving the representation of anthropogenic CO2 emissions in climate models: Impact of a new parameterization for the Community Earth System Model (CESM)"

The reason is that main goal of the paper is not to describe the new parameterization but to evaluate its performance and impact.

Reply: Indeed, the suggested title describes more precisely the aim of the paper. Thanks. The title now reads:

> ***Improving the representation of anthropogenic CO_2 emissions in climate models: impact of a new parameterization for the Community Earth System Model (CESM).***

Referee #3: *Abstract*

"The results show that it is indeed advantageous to model CO_2 emissions and pollutants directly at model grid points rather than using the forcing approach". Please, reword as it is not clear (at this point) what is this forcing approach.

Reply: We rewrote the sentence to make the point clearer.

The text reads:

> *The results show that it is indeed advantageous to model CO_2 emissions and pollutants directly at model grid points rather than using the same mean value globally.*

Referee #3:

Introduction:

The reader would benefit from a more detailed description of the existing approaches to modeling CO2 emissions. What I gather from the paper is the following: there are two basic approaches that models use to account for CO_2 forcing:

. *a) using globally homogenous forcing;*

. *b) using non-homogenous, grid-point specific forcing. This one can be applied in several ways:*

1. using Representative Concentration Pathways (RCPs) that "are not fully- integrated socioeconomic parameterizations, but rather estimates for describing plausible trajectories of human climate change drivers .... They provide simplified accounts of human activities and processes, including population density and economic development, from non-coupled Integrated Assessment Models (IAMs;)" Question: are these parameters location- specific? This is what I am understanding.

2. the proposed here POPEM model being integrated into a fully coupled model. This is similar to RCPs but: uses a coupled model; uses a dynamic model for the prediction of population and emissions.

. *c) Is my understanding correct???*

Reply: Our apologies. We did not make the point clear. It is the other way around: RCPs are used as a surrogate for point-wise estimates. We have clarified that in the revision of the paper [see next comments for more details]

Referee #3:

. *d) If so, I would suggest two possible modifications:*

1. Use some wording or structure as what I've described above

2. Space-permitting, create either a small table or a flow diagram that shows these different levels of sophistication

Reply: We have rewritten the two paragraphs to clarify the differences between RCPs and POPEM approaches. Thanks.

The amended paragraphs now read:

One of the fields most in need of development is the inclusion of co-evolutionary dynamical interactions of the socioeconomic dimension into global models with other Earth system components (Nobre et al., 2010; Robinson et al., 2017; Sarofim and Reilly, 2011). Human activity was a major driver of change in the Earth System in the recent past (Alter et al., 2017; Barnett et al., 2008; Crutzen, 2002), and it now dominates the natural system (Ruth, et al. 2011). However, most global models use basic socioeconomic assumptions about the behavior of societies and are only unidirectionally linked to the biogeophysical part of the Earth system (Müller-Hansen et al., 2017; Smith et al., 2014). The standard way of introducing anthropogenic climate change into ESMs is through Representative Concentration Pathways (RCPs). These are consistent sets of projections involving only radiative forcing *components (van Vuuren et al., 2011), but which represent a step forward from the scenario approach of the last decade (Moss et al., 2010; van Vuuren et al., 2014; van Vuuren and Carter, 2014). However, RCPs are not fully-integrated socioeconomic parameterizations but rather estimates for describing plausible trajectories of human climate change drivers (Moss et al., 2010; Vuuren et al., 2012). They provide simplified accounts of human activities and processes from one-way coupled Integrated Assessment Models (IAMs, Müller-Hansen et al., 2017).*

The use of RCPs is advantageous because they provide a set of pathways that serve to initialize climate models. However, two major problems remain within this approach. Firstly, human activities are not intrinsically embedded into the ESM, impeding sensitivity studies. Secondly, because of the weak coupling of IAMs, they cannot capture the sometimes counterintuitive bidirectional feedback and nonlinearity between the socioeconomic and natural subsystems (Motesharrei et al. 2016; Ruth et al. 2011). Good examples that illustrate the importance of including such bidirectional feedbacks feature in the HANDY model (Motesharrei et al. 2014) which has been used to analyze the key mechanisms behind societal collapses using the predator-prey model.

The RCP approach has been used in climate models because of its low computational cost. However, advances in computational resources now allow to parameterize human-Earth processes in a more detailed way, including the inclusion of population dynamics into the modeling, as in the POPEM (POpulation Parameterization for Earth Models) module (Navarro et al., 2017).

Referee #3: *P. 2, lines 25-30 – It says: "Given the highly non-linear character of the processes involved, it is not unreasonable to assume that location is significant, and the spatial and time distribution of these emissions may affect global climate" – a bit unclear. Might be better to say ", it is not unreasonable to assume that specifying (or accounting for) geographical variability is significant"*

Reply: We modified the expression following your suggestion. The text now reads:

Given the highly non-linear character of the processes involved, it is not unreasonable to assume that accounting for geographical variability is significant, and the spatial and time distribution of these emissions may affect global climate (Alter et al., 2017; Grandey et al., 2016; Guo et al., 2013).

Referee #3: *P. 3, lines 2-4: "The aim of this paper is to show that this grid point scale modeling of anthropogenic CO2 emissions (and other pollutants) represents an improvement, and that two important variables, namely global precipitation distribution and surface temperature, are not negatively affected by this more-detailed approach." While this is true I believe this is a rather weak statement regarding the benefits of using POPEM-type parameterization of emissions forecasting. I believe the authors are in a position to make a stronger statement, namely: including the POPEM dynamical forecasting approach that accounts for the spatial and temporal variability of the emission sources, leads to better representation of the geographical variability of the precipitation.*

Reply: We rewrote the las part of the paragraph to include your suggestion.

The text now reads:

> *The aim of this paper is to show that this grid point scale modeling of anthropogenic CO_2 emissions (and other pollutants) represents an improvement over simpler approaches, and leads to better representation of the geographical variability of precipitation.*

Referee #3: *Space-permitting, I would suggest that the* ***Introduction*** *ends with a short description of the outline for the following presentation. Something like: " the following sections outline: the unique features of POPEM; the validation of the POPEM stand-alone performance; the framework for evaluating the impact of POPEM – incorporation into CESM and framework for testing; the comparison between a control run and a POPEM-specific one: evaluating the differences between the two; evaluating how each compares to observations; discussions; summary and conclusions;" This would give the reader a clear structure of the paper to follow and will make it easier to highlight the contributions of the paper.*

Reply: Thanks for the suggestion. We added a new paragraph with a short description of the outline.

The new paragraph reads:

> *The paper is organized as follows: in section 2, we present the validation of the POPEM standalone mode and set the framework for evaluating the impact of POPEM parameterization –its incorporation into the CESM and the testing framework; in section 3, we compare the outputs of CONTROL and POPEM runs and see how they compare with observations. In the conclusion and future work section, we highlight the importance of the dynamical*

modeling of anthropogenic emissions at grid point scale to better represent the socioeconomic parameters in the CESM model and improve precipitation estimates.

Referee #3:

Section 2.2

currently there are sections 2.2 and 2.2.1 but not 2.2.2 or more. It seems that there is no need for 2.2.1. If there is no 2.2.2. I would suggest the following: "2.2 POPEM specifics and validation", followed by "2.2.1 POPEM parameterization model overview: Unique features" and "2.2.2 POPEM trend verification". Of course, this is just a suggestion.

Reply: Thanks for the suggestion. We rewrite subsection titles and numbers to have a clearer structure.

Now, subsections titles are:

2.2 POPEM specifics and standalone validation

2.2.1 POPEM parameterization model overview

2.2.2 POPEM trend verification

Referee #3: *P. 6, lines 8-9 – "Our control case used global CO2 concentration parameters (standard procedure in ESMs), while the POPEM case used geographically-distributed CO2 emissions data" - is the control using homogeneous CO2 concentrations? I am pretty sure this is the case but it might be better to say it this way.*

Reply: [already discussed above] We have replaced the word '**global'** with the word '**homogeneous'** to make it clearer.

Text now reads:

Our control case used ***homogeneous*** *CO_2 concentration parameters (standard procedure in ESMs), while the POPEM case used geographically-distributed CO_2 emissions data.*

Referee #3:

Section 3.1

P.7, line 23 – it appears that figures 6C, 6D, 8C and 8D are referenced before figures 4 and 5 (and the figure 8 is referenced before Fig.7). This should not be the case. The figures should be referenced in order. However, it seems that this is because the current order of the discussions here might need to be modified. Below is what I mean.

a) Maybe the order should be: 1. Test for "no harm" – figures 6C-6D and 8C-8D show that. 2. Compare the CONTROL to the POPEM simulations to see where exactly they differ. 3. Compare both the CONTROL and the POPEM CESM simulations to the observations, looking at regional distributions. The comparison in steps 2 and 3 brings up the impact of the POPEM geographically-aware CO2 emissions on the geographical distribution of the precipitation, highlighting the positive impact POPEM has (especially in step3).

b) Steps 2 and 3 could be switched – depending on what the authors think.

c) I want to point out that the proposed change in the order of the presentation is just a suggestion for the authors to consider.

Reply: Thanks for the suggestion. We have restructured the order of the figures to make it clear.

Referee #3: *P.8, lines 2-3: "It is clear from the figure that POPEM does alter the spatial pattern of precipitation and exerts a definite effect on the climate pattern, as the module reduces the otherwise exaggerated ITCZ precipitation in the Southern Hemisphere (South East Asia and Australia)." Do you have a reference that it was exaggerated?? If so, then this is a very strong point that needs to be emphasized. Also, do you mean Fig. 4 or Fig. 5? Please, specify.*

Reply: The double ITCZ bias is a persistent problem in most climate models. It has been reported by several authors (Mechoso, 1995; Terray, 1997; Lin 2007) and the causes of this bias are still unclear (Li and Xie, 2014). In the Southern Hemisphere, climate models produce an excess of precipitation in the band 10S-15S when compared with satellite observations (Hwang and Frierson, 2012). We have added a few citations to highlight the importance of this issue.

Additionally, we made a new figure (Figure 9) to clarify the improvements of POPEM in the double ITCZ bias [see the next reply].

The paragraph now reads:

> *It is clear from Figures 5A and 6A that POPEM does alter the spatial pattern of precipitation and exerts a definite effect on the climate pattern, as the module reduces the otherwise exaggerated ITCZ precipitation in the Southern Hemisphere reported by several authors (Hwang and Frierson, 2013; Lin and Xie 2014).*

Referee #3: P. 8, lines 7-8: "There are also important differences in precipitation in the 30N-30S band. Here POPEM reduces model bias, especially in the Southern Hemisphere and on the Tibetan Plateau." How do we know that the model bias is reduced?

Reply: We have now explained this point in full in the section 3.2 and also made a new figure to clarify the point (Figure 9).

Figure 9A shows monthly precipitation for the area affected by the double ITCZ bias in the Southern Hemisphere (20S-0, 80E-100W). It is clear from this figure that POPEM yields more realistic representation of precipitation especially in the driest months (June-October). Figures 9B and 9C show the annual cycle of rainfall over the Australia Top End region and over the Tibetan Plateau, respectively. In both instances there is a usual bias in the original CESM. We have noted that despite POPEM obtaining slightly better results, both CONTROL and POPEM still have difficulties to estimate the precipitation of the rainiest months.

The paragraph now reads:

Another important benefit of POPEM is the reduction of the double ITCZ bias in the Southern Hemisphere. Although a small change can be inferred from Figure 7A-B, the improvement is buried in the annual mean precipitation maps. Figure 9A shows that the POPEM results are closer to observations of the intra-annual variability of precipitation, especially for the driest months (June-October).

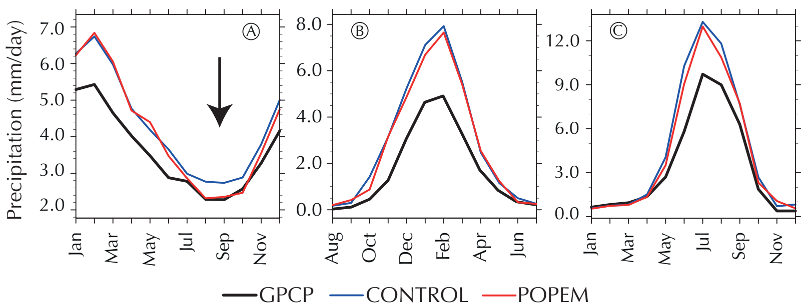


Figure 9: Monthly precipitation (1980-1999) based on GPCP, CONTROL and POPEM for three of the regions with important biases in CESM. (A) shows precipitation for the area affected by the double-ITCZ bias in the Southern Hemisphere (20S-0, 80E-100W); (B) for Australia Top End (30S-10S, 128E-140E); and (C) for the Tibetan Plateau (22N-32N, 78W-92W). The black line represents observations (GPCP), the blue line is the CONTROL case, and the red line is the POPEM case. Units are in mm/day. The arrow indicates the improvement of the POPEM model.

The figure also shows slight improvements for another two typical biases seen in CESM, namely the excess precipitation in the Tibetan Plateau (Chen and Frauenfeld, 2014; Su et al., 2013; Figure 9C) and the bias in some areas affected by the Asian-Australian monsoon (AAM), such as the Australia Top End (Meehl and Arblaster, 1998; Meehl et al. 2012; Figure 9B).

Referee #3: *P. 8, line 9-10: "On the other hand, POPEM departs from the control simulation in the Asia-Pacific region between 10N-10S." Is that good or bad? How do we know?*

Reply: If we zoom-in on figure 6A (map: CONTROL minus POPEM) it can be seen that POPEM produces more precipitation than CONTROL. That means that the model reinforces the double ITCZ bias in this area, which is not good. We have noted that in the paper.

The text reads now:

On the other hand, POPEM departs from the control simulation in the Asia-Pacific region between 10N-10S. This result reinforces the double ITCZ bias in this area.

Referee #3: *P. 8, line 31 – "(Q1 and Q3 remain between ± 0.4 mm/day)." Please, define Q1 and Q3.*

Reply: Q1 and Q3 mean Quartile 1 and Quartile 3. We now write down the word in full to avoid possible confusion.

The line now reads:

(The first and the third quartiles of the distribution remain between ±0.4 mm/day)

References

Hwang, Y.-T. and Frierson, D. M. W.: Link between the double-Intertropical Convergence Zone problem and cloud biases over the Southern Ocean., Proc. Natl. Acad. Sci. U. S. A., 110(13), 4935–40, doi:10.1073/pnas.1213302110, 2013.

Li, G. and Xie, S. P.: Tropical biases in CMIP5 multimodel ensemble: The excessive equatorial pacific cold tongue and double ITCZ problems, J. Clim., 27(4), 1765–1780, doi:10.1175/JCLI-D-13-00337.1, 2014.

Lin, J. L.: The double-ITCZ problem in IPCC AR4 coupled GCMs: Ocean-atmosphere feedback analysis, J. Clim., 20(18), 4497–4525, doi:10.1175/JCLI4272.1, 2007.

Meehl, G. A. and Arblaster, J. M.: The Asian-Australian monsoon and El Nino-Southern Oscillation in the NCAR climate system model, J. Clim., 11(6), 1356–1385, doi:10.1175/1520-0442(1998)011<1356:TAAMAE>2.0.CO;2, 1998.

Mechoso, C. R., Robertson, A. W., Barth, N., Davey, M. K., Delecluse, P., Gent, P. R., Ineson, S., Kirtman, B., Latif, M., Treut, H. Le, Nagai, T., Neelin, J. D., Philander, S. G. H., Polcher, J., Schopf, P. S., Stockdale, T., Suarez, M. J., Terray, L., Thual, O. and Tribbia, J. J.: The Seasonal Cycle over the Tropical Pacific in Coupled Ocean–Atmosphere General Circulation Models, Mon. Weather Rev., 123(9), 2825–2838, doi:10.1175/1520-0493(1995)123<2825:TSCOTT>2.0.CO;2, 1995.

Meehl, G. A. and Arblaster, J. M.: The Asian-Australian monsoon and El Nino-Southern Oscillation in the NCAR climate system model, J. Clim., 11(6), 1356–1385, doi:10.1175/1520-0442(1998)011<1356:TAAMAE>2.0.CO;2, 1998.

Meehl, G. A., Arblaster, J. M., Caron, J. M., Annamalai, H., Jochum, M., Chakraborty, A. and Murtugudde, R.: Monsoon regimes and processes in CCSM4. Part I: The Asian-Australian monsoon, J. Clim., 25(8), 2583–2608, doi:10.1175/JCLI-D-11-00184.1, 2012.

Terray, L.: Sensitivity of Climate Drift to Atmospheric Physical Parameterizations in a Coupled Ocean–Atmosphere General Circulation Model, J. Clim., 11(7), 1633–1658, doi:10.1175/1520-0442(1998)011<1633:SOCDTA>2.0.CO;2, 1998.

Response to referee #4.

Referee #4: *The authors developed a novel POPEM parameterization and applied it to CESM to enhance the realism of global climate modeling by improving the direct representation of human activities and climate. They argued that modeling CO2 emissions and pollutants directly at model grid points is a better approach. As such, their new approach will help understand the potential effects of localized pollutant emissions on long-term global climate statistics, thus assisting adaptation and mitigation policies.*

The topic is interesting and the approach is provoking.

Reply: Thank you for your positive feedback.

Referee #4: *However, I am not quite convinced by the validation part (Part 3.2). I therefore recommend major revision.*

Reply: We followed your recommendations. We have expanded section 3.2 and added a new subsection; *3.3 Validation against ESPI and ONI indices.* Please, see following comments for a detailed revision of the updates. Hope the changes can solve your concerns.

Referee #4: *First, I cannot find a remarkable improvement using POPEM based on the comparison of precipitation and temperature biases. There are some differences between POPEM and CONTROL but these differences are buried in the large biases in either set.*

Reply: We have made clearer in the paper that we do not claim to solve the problem of homogenous emissions versus point-wise estimates. We did not state that our contribution produces a remarkable improvement. What we have achieved by now is far more modest: we have shown that including our more-realistic forcings preserves the model ability to produce realistic fields. Nonetheless, some improvements can be seen (we have included additional figures to illustrate the improvements). We agree that the improvements are limited, but given the small model sensitivity to this forcing (the logic of RCP85 is to somehow 'exaggerate' the emissions to increase the signal), one cannot expect major changes. In other words, the actual signal is too faint to be affected by a more realistic emission pattern. Indeed, the reason for having a distributed method is to be able to evaluate 'what-if' scenarios (i.e. what happens if China cuts off emissions, or the like). We have added a paragraph at the end of the section 3.1 to explain why the approach is valuable in spite of the marginal improvements compared with validation data.

As referee #5 says, we also believe that the use of local population projections to project emissions at each grid point is novel, and is advantageous to the current practice of using global emissions projections to drive ESMs.

The added paragraph reads:

> *Potential applications of POPEM include not only sensitivity analyses of local CO_2 emissions policies, but also the added feature of performing tests for 'what-if' scenarios. One interesting example would be the climate response under the hypothesis that China and India –the most populated countries in the world- reach US CO_2 per capita emissions rates. Another 'what-if' scenario would be the climate response of an increasingly urbanized world. In both cases, POPEM provides a flexible framework for testing the alternative hypotheses.*

Referee #4: *It is true that observations have uncertainties and a new parameterization does not have to improve the model performance in every aspect. Nevertheless, could the authors show some improvements more robust than the current ones (precipitation and temperature) for validation? Maybe TOA radiation balance, ENSO index, Arctic sea ice, etc?*

Reply: We agree that the analysis of Artic sea ice response would be a good addition. Unfortunately, sea ice was not a focus of our research when we ran the simulations and now it is too late to do so. Same about TOA. However, in order to satisfy this requirement, we have included two additional validation metrics using two ENSO indices: namely the ENSO Precipitation Index (ESPI) and the Oceanic el Niño Index (ONI).

We have chosen the ESPI index, which estimates the gradient of the anomalies across the Pacific basin (Curtis and Adler, 2000). It compares well with SST-and pressure-based indices and is widely used by the scientific community (Figure 13 now). The Oceanic el Niño Index is a SST index developed by NOAA as a principal measure for monitoring, assessing and predicting ENSO (Kouski and Higgins, 2007).

We have made two new figures and added a table: Figure 13 for ESPI index, El Niño (EI) and La Niña (LI), and Table 1 and Figure 14 for ONI.

The new section reads as follows:

3.3 Validation against ESPI and ONI indices

The El Niño-Southern Oscillation (ENSO) is the most dominant inter-annual climate variation in the tropics. It occurs when seasonally averaged sea surface temperature anomalies in the eastern Pacific Ocean exceed a given threshold and cause a shift in the atmospheric circulation (Trenberth 1997). Historically, the definition of ENSO does not include precipitation because of the limitations of stations (Ropelewski and Halpert, 1987), but recent work with satellites has confirmed that this phenomenon is a major driver of global precipitation variability (Haddad et al., 2004).

A major advantage of satellite-derived precipitation indices over more conventional ones is the description of the strength and position of the Walker circulation (Curtis and Adler, 2000). Under that assumption, Curtis and Adler (2000) derived three satellite-based precipitation indices: the ENSO precipitation index (ESPI); El Niño index (EI); and La Niña index (LI). Precipitation anomalies are averaged over areas of the Equatorial Pacific and Maritime Continent -where the strongest precipitation anomalies associated with ENSO are found- to construct differences or basin-wide gradients (Curtis, 2008).

Figure 13 shows a comparison of GPCP, CONTROL, and POPEM for the ESPI, EI and LI indices.

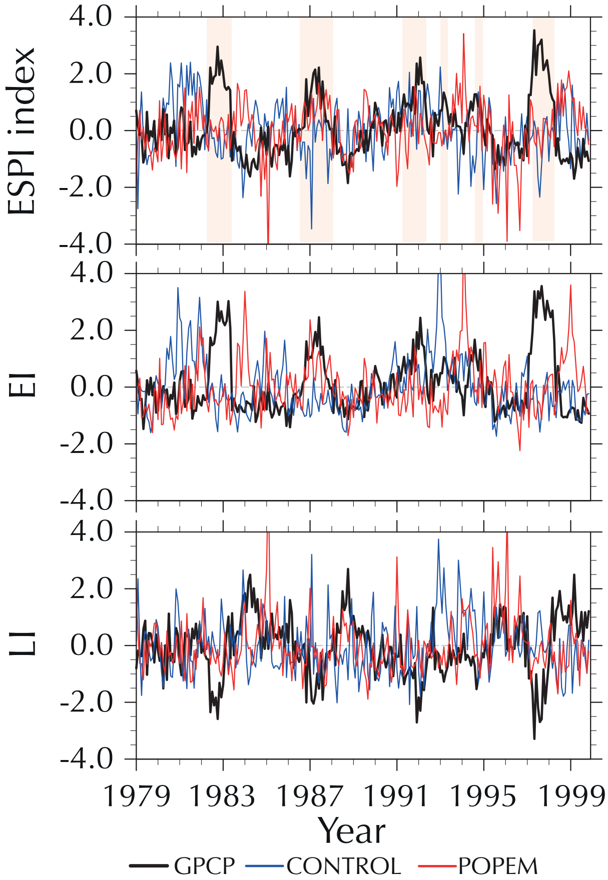


Figure 13: Time-series of precipitation anomalies for the ENSO region after Curtis and Adler (2000). (Top) ENSO Precipitation Index (ESPI); (Middle) El Niño Index (EI); and (Bottom) La Niña Index (LI). The Black line shows GPCP data, the blue line is the CONTROL case, and the red line is the POPEM case. Orange shading denotes El Niño years defined as consecutive months (minimum 3) with NIÑO3.4 sea surface temperature anomalies (5N–5S, 170–120W) greater than +0.5° C.

Unfortunately, CONTROL and POPEM cases have difficulty simulating the precipitation patterns associated with ENSO. Figure 13 shows that bias increases in 82-83 and 97-98 El Niño years. The same bias emerges when comparing the EI and LI indices. In that case, the CESM model produces stronger El Niño/La Niña events than the observed data. Consequently, we can consider that CESM is unable to obtain a precise estimate of precipitation patterns, suggesting that current climate models are far from generating realistic simulations of the precipitation field (Dai, 2006).

Another widely used ENSO index is the Oceanic Niño Index (hereafter ONI). ONI was developed by the NOAA Climate Prediction Center (CPC) as the principal means for monitoring, assessing and predicting ENSO (Kousky and Higgins, 2007). This index is defined as 3-month running-mean values of SST departures from the average in the Niño-3.4 region. It is computed from a set of homogeneous historical SST analyses (Kousky and Higgins, 2007, Smith et al. 2002).

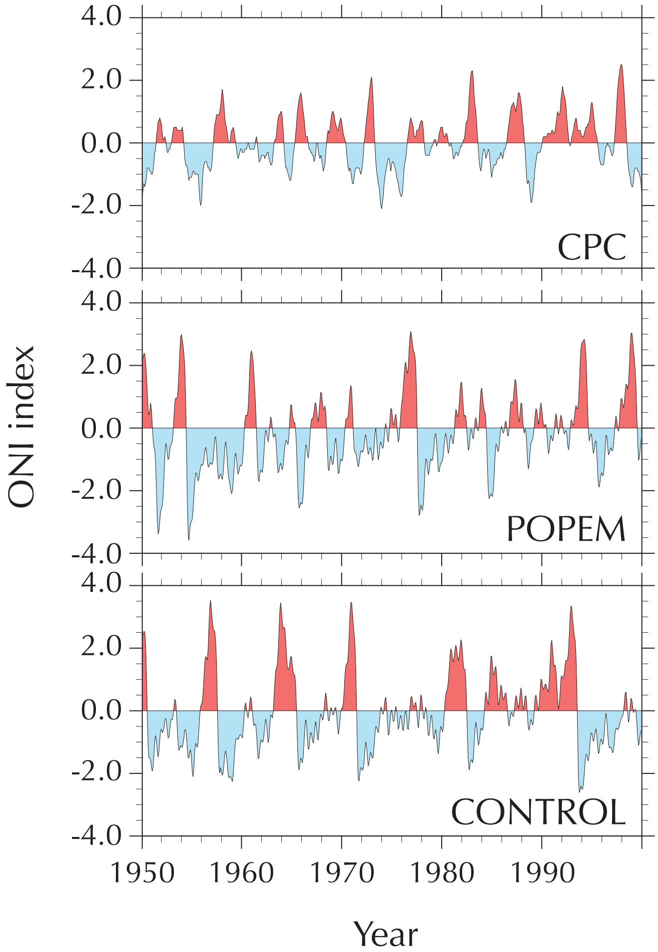


Figure 14: Comparison of the Oceanic el Niño Index (ONI) for CPC (top), POPEM (middle), and CONTROL (bottom) cases. El Niño and La Niña are defined according to Kousky and Higgins (2007): 3-month running mean with anomalies greater than +0.5ºC (or-0.5ºC) for at least five consecutive months in NIÑO3.4 region. The base period for computing SST departures is 1971–1999.

Figure 14 compares the ONI index for CPC, POPEM and CONTROL cases. It is clear from the figure, that POPEM produces a more realistic representation of the ENSO, especially if we focus on the 1992-1999 period. POPEM also obtains better results than CONTROL in the number of simulated el Niño events (see Table 1). The improvement is also noticeable in the intensity. The CONTROL case exhibits an overly strong ENSO -a common bias in CESM (Tang et al., 2016)- but POPEM reduces this bias (0.22^o C versus 0.59^o C).

Table 1. Comparison of the ONI index for the period 1950-1999. The table compares the ability of the models to reproduce the number, strength, and duration of el Niño events.

Source	Number of events	Agreement [1]	Disagreement [2]	Intensity Bias $_{avg}$ [3]	Duration $_{avg}$ [4]
CPC	14				10.3
CONTROL	7	33	121	0.59 ^{o} C	19.4
POPEM	10	37	121	0.22^o C	11.4

[1] The number of months that CPC and CESM agree on El Niño.[2] Disagreement defined as the number of months where CPC and CESM obtain opposite results. [3] Intensity: (|CESM ONI| – |CPC ONI|)/number of cases (units in degrees Celsius). [4] Mean duration of El Niño event (in months).

Another important indicator is the mean duration of El Niño events. Table 1 shows that POPEM obtains better results according to observations (11 months in CPC, 10 months in POPEM, and 19 months in CONTROL).

Referee #4: *Actually, I am somewhat interested in the Arctic sea change. It is known that climate models (like CESM CONTRL) cannot capture a rapid observed decline of Arctic sea ice during recent decades. In Fig. 5(B), POPEM is colder than CONTROL over the Barents Sea area. Will this mean that Arctic sea ice decline in POPEM is even slower than that in CONTROL?*

Reply: It's true that the POPEM parameterization produces colder temperatures in that area and that might reinforce the bias of a slower Artic sea ice decline. Unfortunately, we can't contrast this hypothesis because we did not keep the sea ice outputs for our simulations. Sorry about that.

The bias is less evident when confronted with GISTEMP annual mean anomalies for that area. It is seen from the Figure 11 (top) that CONTROL and POPEM cases have a similar margin error. In other words, the original CESM model is not really good in capturing this feature. Our approach slightly improves the situation in some cases (Bering Sea from 1975 to 1990, Figure 11 (middle)) but we cannot expect a major overall improvement.

We have added a paragraph and a figure to clarify this point.

The text now reads:

The bias is also reproduced when compared with temperature anomalies for a specific region. Thus, for instance, CESM gives poor scores in the Barents Sea area (Figure 11; top) while POPEM obtains better results for the Bering Sea, especially in the Russian part (Figure 11; middle). Here, POPEM gives more realistic values for the period 1970-1998 but, even with the improvement, the model still overestimates the temperature anomaly.

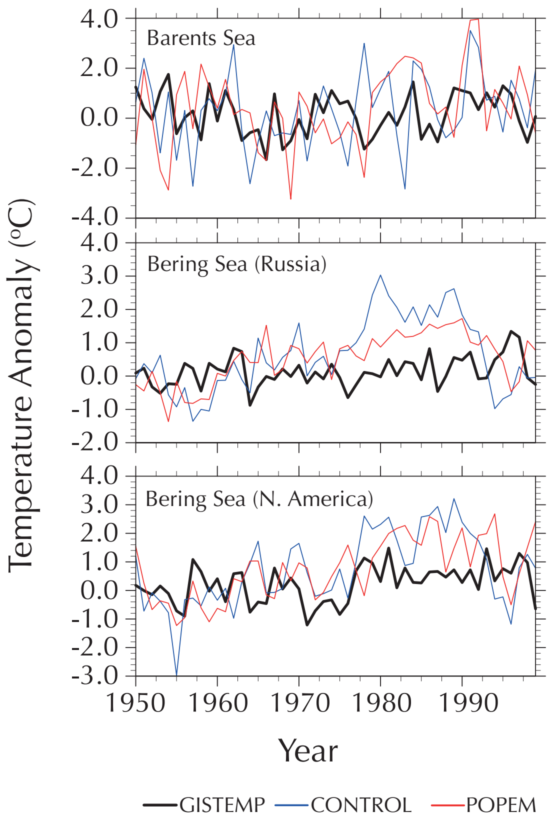


***Figure 11**: A comparison of the annual mean surface temperature anomaly between GISTEMP, CONTROL and POPEM from 1950 to 1999. (Top) represents the Barents Sea (68N-80N, 19E-68E); (middle) Russian part of the Bering Sea (50N-65N, 150E-180E); and (bottom) American part of the Bering Sea (50N-75N, 140W-180W). The black line represents observational data (GISTEMP), the blue line is the CONTROL case, and the red is the POPEM case. Anomaly was referenced to 1951-1980 period.*

We also calculated the temperature anomalies with monthly data (attached as a supplementary material). However, the noise is high and it is difficult to distinguish any clear pattern other than the consistency between the series. Only in Figure EXT2(top) we see that POPEM more frequently yields extreme values.

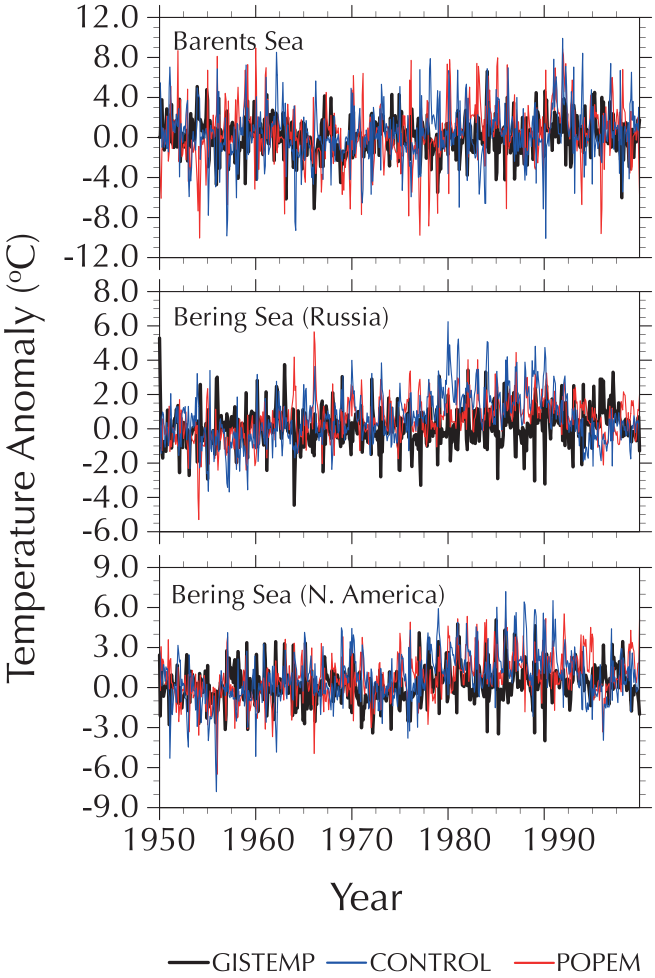


Figure EXT2: The same as Figure 11 but using monthly means.

Referee #4: *Besides, to be consistent with GPCP, the authors may want to use a globally (land+ocean) covered temperature dataset GISTEMP (https://data.giss.nasa.gov/gistemp/) to examine temperature bias.*

Reply: Thanks for the suggestion. As you seen in the previous comment we included GISTEMP in several figures and also made a brief description of the source in section 2.4.3.

The new subsection reads:

> ***2.4.3 GISTEMP data set***
>
> *NASA's GISTEMP (GISS Surface Temperature Analysis) is a global surface temperature change dataset (Hansen and Lebedeff, 1987; see Hansen et al. 2010 for an updated version). It combines land and ocean surface temperatures to create monthly temperature anomalies at 2^o x 2^o degrees of spatial resolution. The use of anomalies reduces the estimation error in those places with incomplete spatial and temporal coverage (Hansen and Lebedeff, 1987). The anomalies are calculated over a fixed base period (1951-1980) that makes the anomalies consistent over long periods of time.*
>
> *The first version was originally conceived only for land areas (Hansen and Lebedeff, 1987) but in 1996 marine surface temperatures were added (Hansen et al., 1996). The updated version of GISTEMP includes satellite-observed nightlights to identify stations located in extreme darkness and adjust temperature trends of urban stations for non-climatic factors (Hansen et al. 2010). Just like CRUTS, GISTEMP is commonly used to validate climate models because of its coverage and confidence levels (Baker and Taylor, 2016; Brown et al., 2015; Neely et al., 2016, Peng et al., 2015).*

Additionally, we used GISTEMP to analyze temperature anomalies for regional (previous comment; Figure 11) and global scales (Figure 12).

The results of Figure 12 were discussed in the section 3.2 of the manuscript:

The new paragraph reads as follows:

> *If we focus on global temperature anomalies, CESM simulations are able to reproduce the progressive increase in the temperature anomaly (Figure 12; top). However, the CONTROL case simulates a sharp drop at the end of the period (1990-1999), while POPEM portrays this change as smooth, in agreement with the observations.*

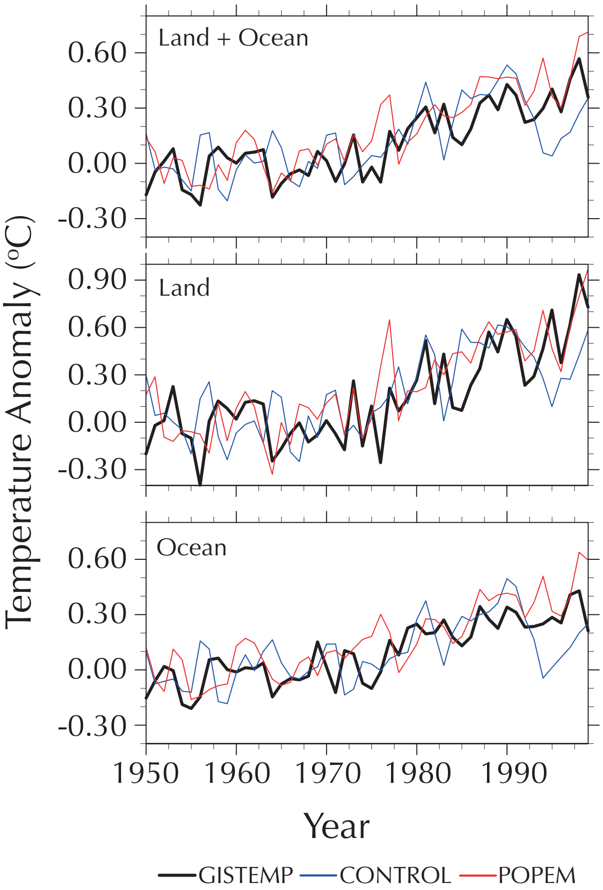


Figure 12: A comparison of the global annual mean surface temperature anomaly between GISTEMP, CONTROL, and POPEM from 1950 to 1999. (Top) global; (middle) land; and (bottom) ocean. The black line represents observational data (GISTEMP), the blue line is the CONTROL case, and the red is the POPEM case. Anomaly was referenced to 1951-1980 period.

The differences between CONTROL and POPEM are better demonstrated when comparing land and ocean separately (Figure 12; middle and bottom). While the temperature anomalies for land are quite similar in both cases, POPEM provides a better representation of the ocean tendency from 1992 onwards, and that translates to an overall improvement (Figure 12, top).

We also computed monthly mean temperature anomalies. However, is difficult to appreciate the differences between models, especially for cases A and B. The figure is therefore included as a supplementary material.

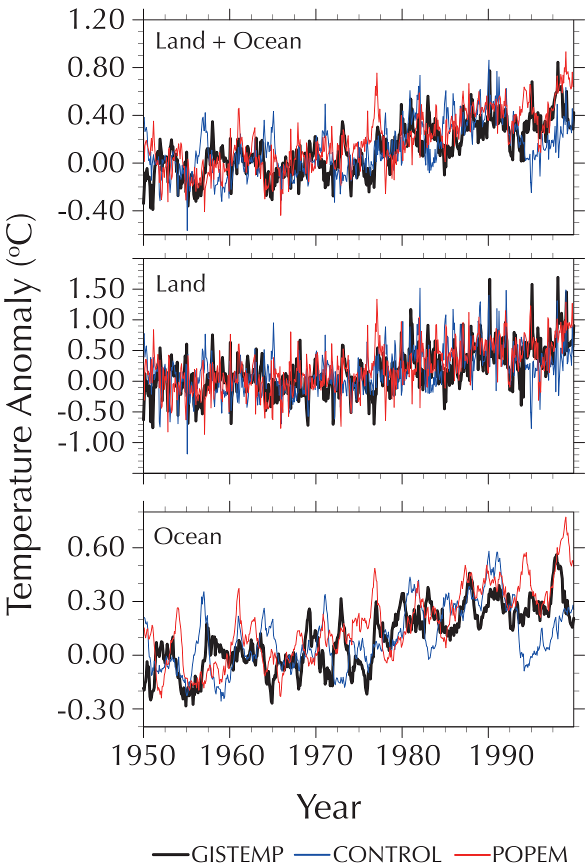


Figure EXT3: Same as Figure 12 but for monthly mean temperature anomaly. The main tendency is consistent albeit differences exists. Thus for instance the POPEM model clearly improves over CONTROL from 1992 onwards.

References

Curtis, S., Adler, R., Curtis, S. and Adler, R.: ENSO Indices Based on Patterns of Satellite-Derived Precipitation, J. Clim., 13(15), 2786–2793, doi:10.1175/1520-0442(2000)013<2786:EIBOPO>2.0.CO;2, 2000.

Kousky, V. E. and Higgins, R. W.: An Alert Classification System for Monitoring and Assessing the ENSO Cycle, Weather Forecast., 22(2), 353–371, doi:10.1175/WAF987.1, 2007.

Response to referee #5. S. Motesharrei

Referee #5:

Review of the manuscript "Improving the representation of anthropogenic CO_2 emissions in climate models: a new parameterization for the Community Earth System Model (CESM)" by Andrés Navarro, Raúl Moreno, and Francisco J. Tapiador, submitted to the Journal Earth System Dynamics, European Geosciences Union (EGU).

Decision:

Because of the importance of the topic, I would recommend the publication of this manuscript after major revisions in the presentation of the work as described below.

Reply: Many thanks for your positive feedback. Please, see following comments for a detailed revision of the updates.

Referee #5:

General comments:

The authors acknowledge (but not completely clearly) a major shortcoming of the Earth System Models (ESMs) and Integrated Assessment Models (IAMs).

Even though the Human System has become the dominant driver of most components of the Earth System since about 1750, and especially since about 1950, IAMs use independent, exogenous projections of the Human System (HS) variables in order to drive ESMs to create future projections. Not including essential bidirectional feedbacks between ES and HS can lead to missing important dynamics that is critical to the sustainably of our planet and people. This problem is discussed in detail in the "Modeling Sustainability" paper by Motesharrei et al. [2016]:

Motesharrei, Safa, Jorge Rivas, Eugenia Kalnay, Ghassem R. Asrar, Antonio J. Busalacchi, Robert F. Cahalan, Mark A. Cane, et al. "Modeling Sustainability: Population, Inequality, Consumption, and Bidirectional Coupling of the Earth and Human Systems." National Science Review 3, no. 4 (December 11, 2016): 470–494. https://doi.org/10.1093/nsr/nww081.

Reply: Many thanks for giving us the opportunity to expand this point in the paper. Also, thanks for the reference, which reinforce our point. We have used it to expand the issue in the revised version of the manuscript.

We extended the third paragraph of the introduction section to explain the point in details. The text now reads:

One of the fields most in need of development is the inclusion of co-evolutionary dynamical interactions of the socioeconomic dimension into global models with other Earth system components (Nobre et al., 2010; Robinson et al., 2017; Sarofim and Reilly, 2011). Human activity was a major driver of change in the Earth System in the recent past (Alter et al., 2017; Barnett et al., 2008; Crutzen, 2002), and it now dominates the natural system (Ruth, et al. 2011). However, most global models use basic socioeconomic assumptions about the behavior of societies and are only unidirectionally linked to the biogeophysical part of the Earth system (Müller-Hansen et al., 2017; Smith et al., 2014). The standard way of introducing anthropogenic climate change into ESMs is through Representative Concentration Pathways (RCPs). These are consistent sets of projections involving only radiative forcing components (van Vuuren et al., 2011), but which represent a step forward from the scenario approach of the last decade (Moss et al., 2010; van Vuuren et al., 2014; van Vuuren and Carter, 2014). However, RCPs are not fully-integrated socioeconomic parameterizations but rather estimates for describing plausible trajectories of human climate change drivers (Moss et al., 2010; Vuuren et al., 2012). They provide simplified accounts of human activities and processes from one-way coupled Integrated Assessment Models (IAMs, Müller-Hansen et al., 2017).

The use of RCPs is advantageous because they provide a set of pathways that serve to initialize climate models. However, two major problems remain within this approach. Firstly, human activities are not intrinsically embedded into the ESM, impeding sensitivity studies. Secondly, because of the weak coupling of IAMs, they cannot capture the sometimes counterintuitive bidirectional feedback and nonlinearity between the socioeconomic and natural subsystems (Motesharrei et al. 2016; Ruth et al. 2011). Good examples that illustrate the importance of including such bidirectional feedbacks feature in the HANDY model (Motesharrei et al. 2014) which has been used to analyze the key mechanisms behind societal collapses using the predator-prey model.

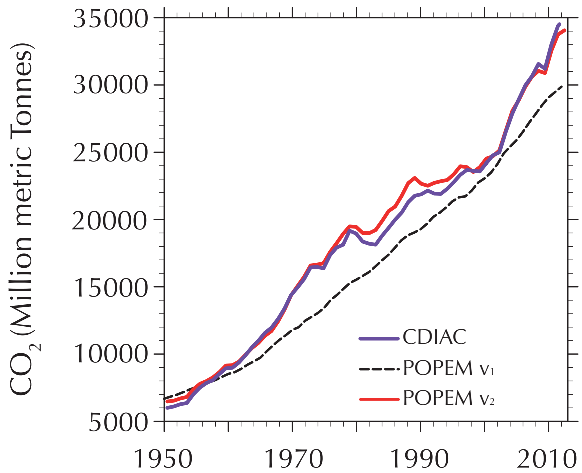
The RCP approach has been used in climate models because of its low computational cost. However, advances in computational resources now allow to parameterize human-Earth processes in a more detailed way, including the inclusion of population dynamics into the modeling, as in the POPEM (POpulation Parameterization for Earth Models) module (Navarro et al., 2017).

***Referee #5:** The manuscript is closely related to a recently published paper by the same team of authors (and, unfortunately, there is much overlap with that already published work):*

Navarro, Andrés, Raúl Moreno, Alfonso Jiménez-Alcázar, and Francisco J. Tapiador. "Coupling Population Dynamics with Earth System Models: The POPEM Model." Environmental Science and Pollution Research, September 16, 2017, 1–12. https://doi.org/10.1007/s11356-017-0127-7.

Reply: Actually, there are major differences with that paper. Navarro et al. (2017) described in detail the **demographic** part of POPEM. In that paper, we were focused on the explanation and validation of the demographic and emission parts at global scale. In contrast, the current paper deals with the coupling of that demographic model with an **Earth System Model**, and compare the model outputs with observational data. That is a completely different history.

Also, the original emissions modeling module has been improved. We included a new figure (EXT1) in the supplementary material to show that. It looks:



EXT1: Comparison of the historical global CO_2 emission estimates for the years 1950–2012. The black line shows the estimates given using POPEM v1, red indicates POPEM v2, and purple indicates CDIAC estimates. Values are given in million of metric tonnes.

We have now limited the potential overlaps to the minimum required for the paper to be self-contained. We have rewritten parts of that section and added a new paragraph in the *2.2.1 POPEM parameterization model overview* subsection to clarify the novelties between successive POPEM versions and how the changes affect the emission estimates and the coupling with the model.

The new paragraph reads:

The demographic/emissions module presented here is an updated version of the demographic module explained in Navarro et. al (2017). The differences between the versions are minimal. They involve better approximation of emissions in highly polluting regions with poor population data, such as China; a better estimate for coastal zones and country limits; and a change in the model time step for more efficient coupling with CESM. The inclusion of these changes results in more accurate emissions estimates when compared with inventories than the previous version did. However, the model is not immune to bias. The most important limit is the degradation of the model outputs when there is increased spatial resolution –resolution of 0.25º and higher–.

Referee #5: *These two papers take a step toward including at least parts of the Human System (human population and emissions) explicitly in the ESMs, however, the somewhat in- accurate presentation of the work (and occasional over-statements) may lead to readers' confusion about the extent and novelty of this work. During my initial reading of the manuscript, I was very impressed by the model and thought that it is a bidirectionally coupled Human System + Earth System Model. (It seems Anonymous Referee 3 has this same impression.) But upon further reading of the manuscript as well as Navarro et al. [2017], I realized that POPEM is essentially a demographic projection model (although it uses dynamic variables for age cohorts) that is used to drive CESM.*

Reply: Sorry if the description of POPEM in the first version of the manuscript was unclear. We have now amended the explanation to avoid the confusion [cf. reply to section (A)].

Referee #5: *By contrast, I believe the use of local population projections to project emissions at each grid point is novel, and is advantageous to the current practice of using global emissions projections to drive ESMs.*

Reply: Thank you for noting this. We believe that this is the central idea of the paper.

Referee #5:

Suggested Revisions:

The other three referees already provide many helpful, important suggestions to improve the manuscript. Here, I outline some additional suggestions to help accurately present the model, its value for the Earth System modeling community, and possible future steps that needs to be taken by the modeling community to make the projections of the "Earth–Human System Models" more realistic.

Reply: Many thanks for your valuable comments to improve the model.

Referee #5: *I do not ask for any changes to model, since such changes would require major effort and could be implemented in future versions.*

Reply: Thanks for your understanding and consideration; really appreciate it.

Referee #5: *(A) Clarify that POPEM is, after all, a demographic projection model. It is true that its 18 age cohorts are dynamic variables, however, they still change based on exogenous fertility and mortality rates*

Reply: Sorry if that was not clear in the first version of the manuscript. We have now extended the first paragraph of *2.2.1 POPEM parameterization model overview* to make it clear. We also redesigned Figure 1 highlighting now the external parameters.

The paragraph now reads:

> *The POPEM module is a demographic projection model coded in FORTRAN that is intended to estimate monthly fossil fuel CO_2 emissions at model grid point scale using population as the input. Due to a lack of actual GHG measurements at appropriate spatial and temporal scales, it is necessary to use a proxy. For this, POPEM employs population, the evolution of which is modeled using external parameters that feed the module.*

Referee #5: *(POPEM does not model Migration, which has become a major driver of population change, especially recently.)*

Reply: Modeling migration flows is an important point that we have taken into account since the very beginning of this project because it is a key element of population change –present and future-. However, there are several restrictions to accuracy estimate migration flows for historical populations at grid cell scale. Firstly, there are two different types of fluxes –short and long distance migrations- that have to be modeled in different ways (Lenormand et al. 2016). Secondly, we must quantify the entering and the exiting population for each cell using a probability rate of migration that is difficult to estimate with the limited migration data (Navarro et al. 2017). Thirdly, it is difficult –but not impossible- to validate a highly-detailed migration model with limited availability of migration data. Fourthly, the computational cost rises dramatically (e.g. *4 types of migration fluxes* x *number of cells* x *age-group* x *number of timesteps*).

Consequently, these sources of uncertainties are greater than the benefits for the period of time and the spatial resolution used in this work.

Referee #5: *These rates are projected into the future using statistical methods such as in the UN Population Projections. Therefore, the projections using POPEM could not be much different from traditional demographic projections, as can be seen from comparisons of POPEM to UN projections in Navarro et al. [2017]. I believe indeed POPEM cannot properly capture demographic change details for some regions and for certain age cohorts. Therefore, the value-added from this 'dynamic' population model is limited, at least from a demographic perspective.*

Reply: We assume that there is room for improvement in the demographic part of the model and it is an important point that we have to develop in the future versions of POPEM. However, the time period that we used here (1950-2000) and the actual spatial resolution offered by POPEM (1^o x 1^o) make model outputs less sensible to the referred biases. We have nonetheless clarified the limitations of the approach in the revised version. [see above the reworked text]

Referee #5: *(B) Because ES and other components of the HS do not feedback onto the demographic variables in POPEM, POPEM will not be able to capture non-trivial dynamics that can arise due to such bidirectional feedbacks [Motesharrei et al., 2016]. For basic examples of how these bidirectional feedbacks (in a minimal model) can lead to surprising behavior, see:*

Motesharrei, Safa, Jorge Rivas, and Eugenia Kalnay. "Human and Nature Dynamics (HANDY): Modeling Inequality and Use of Resources in the Collapse or Sustainability of Societies." Ecological Economics 101 (May 2014): 90–102. https://doi.org/10.1016/j.ecolecon.2014.02.014.

Reply: Firstly, thank you for this crucial reference. We considered that the citation of this work in the first part of the manuscript clarifies how important are the Human-Earth interactions and their feedbacks for models.

Secondly, we agree with you that bidirectional feedbacks between ES and HS are essential to make ESMs more accurate and realistic. The work presented here is just the first step in that direction.

[See the second comment in the discussion to check how we have expanded this point in the revised version of the manuscript.]

***Referee #5:** (C) I strongly recommend adding a schematic diagram at the begging of the paper to show how POPEM interacts with CESM (e.g., variables, parameters, input/output, couplings).*

Reply: Thanks. We have reworked Figure 1 following your recommendations.

Figure 1 now looks:

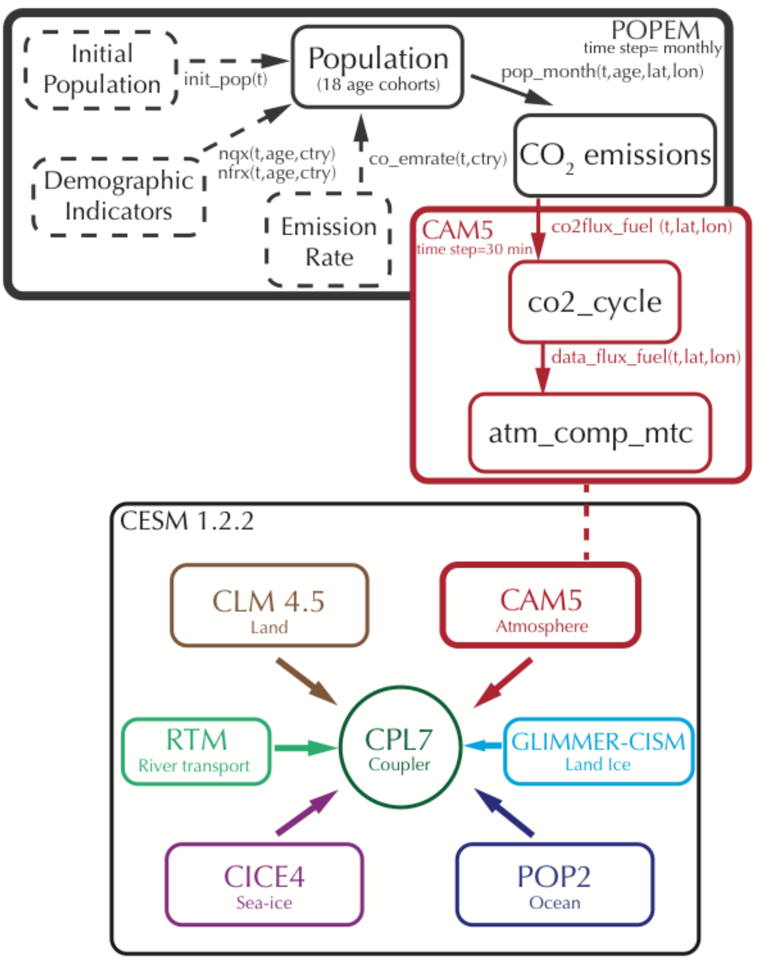


Figure 1: Conceptual schema of the POPEM module coupled with the CAM5 atmosphere module. POPEM requires three input data sets to compute emissions (black dashed rectangles): initial population distribution; demographic parameters (age structure, death, and birth rates); and per capita emission rates by country. POPEM provides a 3D array (time, latitude, longitude) with emissions that are read by the *CO2_cycle* module and passed to the *atm_comp_mct* module which computes the total amount of CO_2 in the atmosphere.

Referee #5: *(D) If POPEM + CESM is indeed the* first *model that calculates emissions at a local scale, as opposed to using global emissions projections, please emphasize that as the novel accomplishment of this research*

Reply: Thanks for the suggestion. We added two sentences in the las part of the first paragraph (section 2.2.1).

The extended version now reads:

> *[...]The idea of using population as proxy is not new, and population density has previously been used to downscale national CO_2 emissions (Andres et al., 1996, 2016). However, these inventories were not dynamical, but instead tied to historical data so it is not possible to use them either to estimate future changes in emissions, or coupled with other components of the model. This change represents an important advance in the way emissions are computed. Thus, POPEM uses a bottom-up approach, where emissions are calculated at cell level on the basis of population dynamics, while global inventories use a top-down approach, which is less flexible when coupled with other components of the ESM.*

Referee #5: *(E) Remove any parts of the manuscript that overlaps with Navarro et al. [2017], and instead refer to specific parts of that publication.*

Reply: We have removed some overlapping text and referred to Navarro et al. 2017. However, there are some elements that it is important to keep in the manuscript for the reasons mentioned at the beginning of this discussion (see reply to third comment). Hope you find the reasons compelling enough to justify our choice.

Referee #5: *(F) Be more careful with the definitions of, and distinctions between, ESMs and IAMs. Navarro et al. [2017] write, for example: "[RCPs] provide simplified versions of human activities and processes, such as population density and economic development, from non-coupled Integrated Assessment Models (IAMs)." It is not true that IAMs are 'non-coupled'; they are indeed one-way coupled.*

Reply: Sorry about that. What we wanted to say here was 'one-way coupled'.

Referee #5: *Then the authors write "researchers in the iESM Project (Collins et al. 2015) developed a global integrated assessment model, the GCAM, to address human impact on climate dynamics, with special emphasis on the representation of the human earth system." GCAM was not developed in the iESM project, but has been in development since 1990s and is one of the leading IAMs. The rest of the description of the sentence is also incorrect. iESM couples land use and agriculture to ES via bidirectional feedbacks.*

Reply: Sorry about that. Perhaps we should have described more precisely that GCAM is the IAM used by the iESM model in that paper. We take note of that for the future.

Referee #5: *(G) In the last section of the manuscript (4), emphasize that dynamic models of various Human System components need to be developed and coupled to ESMs via bidirectional feedbacks in order to produce realistic projections and to capture counterintuitive and unexpected dynamics.*

Reply: Thanks for the suggestion. We added a concluding paragraph in the manuscript.

The new paragraph reads:

> *Although the version of POPEM presented here is already functional, this work is intended to be just the first step in fully coupling socioeconomic dynamics with ESMs. This will include bidirectional feedback between Human and Earth systems and the simulation of societal processes based on the internal dynamics of the model instead of using external sources to make the projections. Only within a coupled global Human-Earth system framework can we produce more realistic representations of the Earth system capturing much of the counterintuitive feedback that is missing from current models (Motesharrei et al. 2016). The success of this approach will depend on the ability of scientists from different research fields to work in an interdisciplinary framework of continuous collaboration.*

Referee #5: *(H) Please go over your citations carefully and make sure that they appear at proper places. Also, the manuscript can benefit from additional important, relevant references. (The bibliography of Motesharrei et al. [2016] could be helpful for this manuscript.)*

Reply: Thank you for the advice and the reference. That excellent review helped us to find new relevant references, such as the previous work done by Matthias Ruth, Eugenia Kalnay and Jorge Rivas. We revised and extended the introduction section and added new citations from the bibliography of Motesharrei et al. (2016). (see the second comment for details on changes in the introduction section).

REFERENCES

Lenormand, M., Bassolas, A. and Ramasco, J. J.: Systematic comparison of trip distribution laws and models, J. Transp. Geogr., 51, 158–169, doi:10.1016/J.JTRANGEO.2015.12.008, 2016.

Navarro, A., Moreno, R., Jiménez-Alcázar, A. and Tapiador, F. J.: Coupling population dynamics with earth system models: the POPEM model, Environ. Sci. Pollut. Res., in press, doi:10.1007/s11356-017-0127-7, 2017.

Improving the representation of anthropogenic CO_2 emissions in climate models: impact of a new parameterization for the Community Earth System Model (CESM)

Andrés Navarro[1], Raúl Moreno[1] and Francisco J. Tapiador[1]

[1]Institute of Environmental Sciences (ICAM), University of Castilla-La Mancha, Toledo, 45004, Spain.

Correspondence to: A. Navarro (Andres.Navarro@uclm.es)

Abstract. ESMs (Earth System Models) are important tools that help scientists understand the complexities of the Earth's climate. Advances in computing power have permitted the development of increasingly complex ESMs and the introduction of better, more accurate parameterizations of processes that are too complex to be described in detail. One of the least well-controlled parameterizations involves human activities and their direct impact at local and regional scales. In order to improve the direct representation of human activities and climate, we have developed a simple, scalable approach that we have named the POPEM module (POpulation Parameterization for Earth Models). This module computes monthly fossil fuel emissions at grid point scale using the modeled population projections. This paper shows how integrating POPEM parameterization into the CESM (Community Earth System Model) enhances the realism of global climate modeling, improving this beyond simpler approaches. The results show that it is indeed advantageous to model CO_2 emissions and pollutants directly at model grid points rather than using the same mean value globally~~forcing approach~~. A major bonus of this approach is the increased capacity to understand the potential effects of localized pollutant emissions on long-term global climate statistics, thus assisting adaptation and mitigation policies.

1 Introduction

The Earth system is a complex interplay of physical, chemical and biological processes that interact in non-linear ways (Ladyman et al., 2013; Lorenz, 1963; Rind, 1999; Williams, 2005). Much effort has been devoted to understanding these complex interactions, and several improvements have been made since the end of the last century.

One of the most important advances in this field has been the use of coupled numerical climate models, dubbed Earth System Models, or ESMs (Edwards, 2011; Flato, 2011; Schellnhuber, 1999). These models aim to simulate the complex interactions of the atmosphere, ocean, land surface, and cryosphere, together with the carbon and nitrogen cycles (Giorgetta et al., 2013; Hurrell et al., 2013; Martin et al., 2011; Schmidt et al., 2014).

However powerful, climate models are far from being perfect (Hargreaves, 2010; Hargreaves and Annan, 2014). Unresolved processes (Williams, 2005), limited computational resources (Shukla et al., 2010; Washington et al., 2009), and model uncertainties (Baumberger et al., 2017; Lahsen, 2005; Steven and Bony, 2013) are ongoing issues that still require attention and further improvement.

One of the fields most in need of development is the inclusion of co-evolutionary dynamical interactions of the socioeconomic dimension into global models with other Earth system components (Nobre et al., 2010; Robinson et al., 2017; Sarofim and Reilly, 2011). Human activity was a major driver of change in the Earth System in the recent past (Alter et al., 2017; Barnett et al., 2008; Crutzen, 2002), and it now dominates the natural system (Ruth, et al. 2011). However, most global models use basic socioeconomic assumptions about the behavior of societies and are only unidirectionally linked to the biogeophysical part of the Earth system (Müller-Hansen et al., 2017; Smith et al., 2014). The standard way of introducing anthropogenic climate change into ESMs is through Representative Concentration Pathways (RCPs). These are consistent sets of projections involving only radiative forcing components (van Vuuren et al., 2011), but which represent a step forward from the scenario approach of the last decade (Moss et al., 2010; van Vuuren et al., 2014; van Vuuren and Carter, 2014). However, RCPs are not fully-integrated socioeconomic parameterizations but rather estimates for describing plausible trajectories of human climate change drivers (Moss et al., 2010; Vuuren et al., 2012). They provide simplified accounts of human activities and processes from one-way coupled Integrated Assessment Models (IAMs, Müller-Hansen et al., 2017).

The use of RCPs is advantageous because they provide a set of pathways that serve to initialize climate models. However, two major problems remain within this approach. Firstly, human activities are not intrinsically embedded into the ESM, impeding sensitivity studies. Secondly, because of the weak coupling of IAMs, they cannot capture the sometimes counterintuitive bidirectional feedback and nonlinearity between the socioeconomic and natural subsystems (Motesharrei et al. 2016; Ruth et al. 2011). Good examples that illustrate the importance of including such bidirectional feedbacks feature in the HANDY model (Motesharrei et al. 2014) which has been used to analyze the key mechanisms behind societal collapses using the predator-prey model.

The RCP approach has been used in climate models because of its low computational cost. However, advances in computational resources now allow to parameterize human-Earth processes in a more detailed way, including the inclusion of population dynamics into the modeling, as in the POPEM (POpulation Parameterization for Earth Models) module (Navarro et al., 2017).

~~One of the fields most in need of development, is the inclusion of co-evolutionary dynamical interactions of the socioeconomic dimension in global models with other Earth system components (Robinson et al., 2017; Sarofim and Reilly, 2011). To date, most global models have used basic socioeconomic assumptions about the behavior of societies and are only unidirectionally linked to the biogeophysical part of the Earth system (Müller-Hansen et al., 2017; Smith et al., 2014). The~~

standard way of introducing anthropogenic climate change into ESMs is through Representative Concentration Pathways (RCPs). These are consistent sets of projections involving only radiative forcing components (van Vuuren et al., 2011) that represent a step forward from the scenario approach of the last decade (Moss et al., 2010; van Vuuren et al., 2014; van Vuuren and Carter, 2014). However, RCPs are not fully-integrated socioeconomic parameterizations, but rather estimates for describing plausible trajectories of human climate change drivers (Moss et al., 2010; Vuuren et al., 2012). They provide simplified accounts of human activities and processes, including population density and economic development, from non-coupled Integrated Assessment Models (IAMs; (Müller-Hansen et al., 2017)).

This general approach is used in climate models because it has a low computational cost. However, advances in computational resources allow us to parameterize human-Earth processes in a more detailed way. This is the main aim of the POPEM (POpulation Parameterization for Earth Models) module (Navarro et al., 2017). Nevertheless, the enormity of this issue means that we have initially restricted the module's integration into ESMs to an exploration of anthropogenic perturbation of the carbon cycle.

One important, but sometimes overlooked process is the direct, regional effect of anthropogenic greenhouse gas (GHG) emissions. Although some GHGs quickly mix in the atmosphere (IPCC, 2014a), their mixing times and lifetimes vary (Archer et al., 2009; Prather, 2007), and localized emissions may produce a transient response in the atmosphere. Given the highly non-linear character of the processes involved, it is not unreasonable to assume that accounting for geographical variability is significant, and the spatial and time distribution of these emissions may affect global climate (Alter et al., 2017; Grandey et al., 2016; Guo et al., 2013).Given the highly non-linear character of the processes involved, it is not unreasonable to assume that location is significant, and the spatial and time distribution of these emissions may affect global climate (Alter et al., 2017; Grandey et al., 2016; Guo et al., 2013). This hypothesis has seldom been investigated, as most current models treat certain GHG emissions as a homogeneously distributed forcing. Thus, for instance, the most typical CESM (Community Earth System Model) simulations prescribe a CO_2 concentration on the assumption that it is well-mixed in the atmosphere (Neale et al., 2012).

This paper describes the results of a 50-year simulation with a simple parameterization of fossil fuel CO_2 emissions at model grid point scale, integrating the POPEM module into the CESM. The aim of this paper is to show that this grid point scale modeling of anthropogenic CO_2 emissions (and other pollutants) represents an improvement, over simpler approaches, and leads to better representation of the geographical variability of precipitation.and that two important variables, namely global precipitation distribution and surface temperature, are not negatively affected by this more-detailed approach.

The purpose of the new modeling is not only to improve precipitation and temperature estimates, but also help understand the carbon cycle feedback, and evaluate the climate sensitivity of the Earth under alternative GHG emission scenarios. While our focus here is anthropogenic CO_2 emissions, the POPEM parameterization can accommodate other GHGs and human-dependent processes in order to advance CESMs towards a comprehensive, fully-coupled modeling of anthropogenic dynamics in the global climate.

The paper is organized as follows: in section 2, we present the validation of the POPEM standalone mode and set the framework for evaluating the impact of POPEM parameterization –its incorporation into the CESM and the testing framework; in section 3, we compare the outputs of CONTROL and POPEM runs and see how they compare with observations. In the conclusion and future work section, we highlight the importance of the dynamical modeling of anthropogenic emissions at grid point scale to better represent the socioeconomic parameters in the CESM model and improve precipitation estimates.

2. Material and methods

2.1 The CESM model

The Community Earth System Model (CESM) is a state-of-the-art ESM and probably the most widely used climate model. It was developed and is maintained by the National Center for Atmospheric Research (NCAR), with contributions from external researchers funded by the U.S. Department of Energy (DOE), the National Aeronautics and Space Administration (NASA), and the National Science Foundation (NSF) (Hurrell et al., 2013). CESM is an ESM comprising a system of multi-geophysical components, which periodically exchange two-dimensional boundary data in the coupler (Craig et al., 2012). It consists of five component models and one central coupler component: the atmosphere model CAM (Community Atmosphere Model; (Tilmes et al., 2015), the ocean model POP (Parallel Ocean Program; (Kerbyson and Jones, 2005); the land model CLM (Community Land Model; (Lawrence et al., 2011); the sea ice model CICE (Community Ice Code; (Hunke and Lipscomb, 2008); and the ice sheet model CISM (Community Ice Sheet Model; (Lipscomb et al., 2013).

CESM –formerly the Community Climate System Model (CCSM)- was conceived as a coupled atmospheric-oceanic circulation model (Boville and Gent, 1998; Collins et al., 2006; Gent et al., 2011; Hurrell et al., 2013; Williamson, 1983). Since the release of the first version, CESM has evolved into a complex Earth System Model now used in different fields. This includes research into atmospheric (Bacmeister et al., 2014; Liu et al., 2012; Yuan et al., 2013), biogeochemical (Lehner et al., 2015; Nevison et al., 2016; Val Martin et al., 2014), and human-induced processes (Huang and Ullrich, 2016; Levis et al., 2012; Oleson et al., 2011), as well as others. The core code of CESM has also been utilized by various research centers for developing their own models (norESM, Bentsen, 2013; CMCC–CESM–NEMO, Fogli and Iovino, 2014; MIT IGSM-CAM, Monier et al., 2013). CESM has been used in many hundreds of peer-reviewed studies to better understand climate variability and climate change (Hurrell et al., 2013; Kay et al., 2015; Sanderson et al., 2017). Simulations performed with CESM have made a significant contribution to international assessments of climate, including those of the Intergovernmental Panel on Climate Change (IPCC) and the CMIP5/6 project (Coupled Model Intercomparison Project Phase 5/6) (Eyring et al., 2016; IPCC, 2014b; Taylor et al., 2012).

A major advantage of CESM over other ESMs is its availability. Some climate models are developed by scientific groups and access to the source code is limited. The CESM source code is free and available to download from the NCAR website. This approach helps improve the model by setting up a framework for collaborative research and makes the model fully auditable. CESM is a good example of a 'full confidence level' model, after Tapiador et al. (2017), where many 'avatars' of the code are routinely run in several independent research centers, and there is an entire community improving the model and reporting on issues and results. However, the model is not immune to bias. One important shortcoming is the poor representation of precipitation in terms of spatial structure, intensity, duration, and frequency (Dai, 2006; Tapiador et al., 2018; Trenberth et al., 2017, Trenberth et al., 2015). Another major bias is the anomalous warm surface temperature in coastal upwelling regions (Davey et al., 2001; Justin Small, 2015; Richter, 2015).

The CESM has been used in many hundreds of peer-reviewed studies to better understand climate variability and climate change (Hurrell et al., 2013; Kay et al., 2015; Sanderson et al., 2017). Simulations performed with CESM have made a significant contribution to international assessments of climate, including those of the Intergovernmental Panel on Climate Change (IPCC) and the CMIP5/6 project (Coupled Model Intercomparison Project Phase 5/6) (Eyring et al., 2016; IPCC, 2014b; Taylor et al., 2012).

2.2 POPEM specifics and standalone validationThe POPEM parameterization module

2.2.1 POPEM parameterization model overview

The POPEM module is a demographic projection model coded in FORTRAN that is intended to estimate monthly fossil fuel CO_2 emissions at model grid point scale using population as the input. Due to a lack of actual GHG measurements at appropriate spatial and temporal scales, it is necessary to use a proxy. For this, POPEM employs population, the evolution of which is modeled using external parameters that feed the moduleThe POPEM module is a set of FORTRAN routines that are intended to estimate monthly fossil fuel CO_2 emissions at model grid point scale using population as the input. Due to a lack of actual GHG measurements at appropriate spatial and temporal scales, it is necessary to use some sort of proxy. For this, POPEM uses a population, whose evolution is modeled. The idea of using population as proxy is not new, and population density has previously been used to downscale national CO_2 emissions (Andres et al., 1996, 2016). However, these inventories were not dynamical, but instead tied to historical data so it is not possible to use them either to estimate future changes in emissions, or coupled with other components of the model. This change represents an important advance in the way emissions are computed. Thus, POPEM uses a bottom-up approach, where emissions are calculated at cell level on the basis of population dynamics, while global inventories use a top-down approach, which is less flexible when coupled with other components of the ESM.

The demographic/emissions module presented here is an updated version of the demographic module explained in Navarro et. al (2017). The differences between the versions are minimal. They involve better approximation of emissions in highly polluting regions with poor population data, such as China; a better estimate for coastal zones and country limits; and a change in the model time step for more efficient coupling with CESM. The inclusion of these changes results in more accurate emissions estimates when compared with inventories than the previous version did. However, the model is not immune to bias. The most important limit is the degradation of the model outputs when there is increased spatial resolution –resolution of 0.25^{o} and higher–.

Detailed information on POPEM and its validation in the demographic realm can be found in (Navarro et al., 2017). In short, from an initial condition, the routine computes the population for each model grid point in a 2D matrix and then calculates fossil fuel CO_2 emissions using per capita emission rates by nations. The process is repeated for each time step (e.g. annually) throughout the simulation period.

Figure 1 about here

As seen in Figure 1, POPEM stores gridded emission data in a 3D array (time, latitude and longitude) to be used by the modified version of the *co2_cycle* module. This module reads emissions data and passes this to the *atm_comp_mct,* which calculates the total amount of CO_2 emissions from different sources (land, ocean and fossil fuel).

2.2.21 POPEM trend verification Population trend verification

Prior to coupling POPEM with CESM we performed several tests to evaluate its ability to reproduce historical population trends and CO_2 emissions. To do this, we ran the module in standalone mode. In a first test, we ran a short simulation (1950-2013) and compared the emissions data with a standard emissions inventory (CDIAC). In a second test, POPEM was run for 70 years (1950-2020) and population estimates were validated against the UN (United Nations) population statistics database for those years when data was available.

Figure 2 about here

As shown in Figure 2, POPEM is capable of satisfactorily simulating the dynamics of the population. Comparison with UN data shows good agreement. However, POPEM presents slight differences from the reference data in some regions. Several of these discrepancies can be explained by the initial model conditions; POPEM uses the same age distribution inside each grid cell to initiate the model (only for the first time-step). This distribution is based on the global average age structure. Consequently, the model overestimates the population in those regions with a more elderly age structure, i.e., Europe and North America, and underestimates areas with younger populations, i.e., Latin America and Asia.

These disparities in population counts have a diverse effect on the outputs in terms of GHG emissions. Thus, for example, the bias in Europe seems to be more important than the bias in Latin America and Oceania. Two principal reasons could explain this: population size, as Europe has a larger population than Oceania, so there is greater bias in the CO_2 emissions estimation; and the per capita emissions rate, as Latin American countries have lower per capita emissions rates than European nations.

It is worth noting here that the POPEM outputs in Figure 2 are clearly non-linear and thus not trivially derived from simply extrapolating population. The North American estimate of CO_2 emissions (second row from the bottom) clearly shows the added value introduced by the model.

Figure 3 shows how POPEM distributes CO_2 emissions for different years in the recent past. In 1950, the majority of emissions tended to be concentrated in the USA and Europe, while in 2000, China, the USA and India were the most polluting countries. This is consistent with the literature: POPEM's estimates generally agree with the emissions maps for the recent past (Andres et al., 1996; Boden et al., 2017; Oda et al., 2018; Rayner et al., 2010), as well as with regional studies on CO_2 emissions (Gately et al., 2013; Gurney et al., 2009).

Figure 3 about here

The regionalized distribution of emissions depicted in Figure 3 represents a vast improvement over the standard procedure of using globally-averaged emissions. Even accounting for rapid mixing of GHGs gases, transient effects are likely to appear given the hemispheric contrast and regional differences in the emissions. The differences in Asia are illustrative of the economic changes in the recent past and the exponential pace of industrialization in that region.

2.3 CESM experimental setup

The CESM used in this work is based on version 1.2.2 (http://www.cesm.ucar.edu/models/). This set includes active components for the atmosphere, land, ocean, and sea ice, all coupled by a flux coupler. The latest atmospheric module CAM5 (Neale et al., 2012) is used to introduce more accurate modeling of atmospheric physics. Additionally, the carbon cycle module is included in CESM's atmosphere, land, and ocean components (Lindsay et al., 2014).

We ran an experiment at 1.9^{o} degrees of spatial resolution for the period 1950-2000. Two simulations were performed to analyze the effects of the regionalized emissions (Figure 3) on the CESM. Our control case used global CO_2 concentration parameters (standard procedure in ESMs), while the POPEM case used geographically-distributed CO_2 emissions data. In the latter, the POPEM module was coupled with the atmospheric CO_2 flux routine to provide monthly gridded CO_2 emissions. The gridded data was used at each time step by the atmospheric routine. Apart from this change, both simulations were identical in order to identify the effects (if any) of the POPEM parameterization.

2.4 Validation data

2.4.1 GPCP data set

Precipitation is one of the key elements for balancing the energy budget, and one of the most challenging aspects of climate modeling. Hence, high quality estimates of precipitation distribution, amount and intensity are essential (Hou et al., 2014; Kidd et al., 2017; Xie and Arkin, 1997). While there are many sources of precipitation data to be used as a reference (see (Tapiador et al., 2012) for a review), only a few qualify as 'full confidence level validation data' (Tapiador et al., 2017).

The Global Precipitation Climatology Project GPCP (Adler et al., 2016) has several products suitable for validating climate models. GPCP-Monthly is one of the most popular precipitation data sets for climate variability studies. It combines data from rain gauge stations and satellite observations to estimate monthly rainfall on a 2.5-degree global grid from 1979 to the present. The careful combination of satellite-based rainfall estimates results in the most complete analysis of rainfall available to date over the global oceans, and adds necessary spatial detail to rainfall analyses over land. Due to its relevance and global coverage, it has been widely used for validating precipitation in climate models (Li and Xie, 2014; Pincus et al., 2008; Stanfield et al., 2016; Tapiador, 2010).

2.4.2 CRU data set

Global surface temperature data sets are an essential resource for monitoring and understanding climate variability and climate change. One of the most commonly used data sets is produced by The Climate Research Unit at the University of

East Anglia (CRU). This group produces a high-resolution gridded climate dataset for land-only areas, the Climate Research Unit Timeseries (CRUTS) (Harris et al., 2014). CRUTS contains monthly time series of ten climate variables, including surface temperature. The data set is derived from monthly observations at meteorological stations. Station anomalies are interpolated into 0.5° latitude/longitude grid cells covering the global land surface and combined with existing climatology data to obtain absolute monthly values (New et al., 1999, 2000). It is commonly used in the validation of climate models because of its confidence levels, together with temporal and spatial coverage, and the fact it compiles station data from multiple variables from numerous data sources into a consistent format (Christensen and Boberg, 2012; Hao et al., 2013; Liu et al., 2014; Nasrollahi et al., 2015).

2.4.3 GISTEMP data set

NASA's GISTEMP (GISS Surface Temperature Analysis) is a global surface temperature change dataset (Hansen and Lebedeff, 1987; see Hansen et al. 2010 for an updated version). It combines land and ocean surface temperatures to create monthly temperature anomalies at 2° x 2° degrees of spatial resolution. The use of anomalies reduces the estimation error in those places with incomplete spatial and temporal coverage (Hansen and Lebedeff, 1987). The anomalies are calculated over a fixed base period (1951-1980) that makes the anomalies consistent over long periods of time.

The first version was originally conceived only for land areas (Hansen and Lebedeff, 1987) but in 1996 marine surface temperatures were added (Hansen et al., 1996). The updated version of GISTEMP includes satellite-observed nightlights to identify stations located in extreme darkness and adjust temperature trends of urban stations for non-climatic factors (Hansen et al. 2010). Just like CRUTS, GISTEMP is commonly used to validate climate models because of its coverage and confidence levels (Baker and Taylor, 2016; Brown et al., 2015; Neely et al., 2016, Peng et al., 2015).

3. Results and discussion

3.1 Comparison~~s~~ between the CONTROL and POPEM runs

It is worth stressing that a parameterization which performs well when tested for the variable it models does not necessarily translate into an overall improvement of the other variables in the model. An accepted practice in climate modeling is to tune ESMs by adjusting some parameters to achieve a better agreement with observations (Hourdin et al., 2017; Mauritsen et al., 2012). These adjustments to specific targets may, however, decrease the model's overall performance (Hourdin et al., 2017), and give poor scores for variables other than those tuned. Thus, for example, if a model is biased with respect to aerosol concentrations or humidity, then improved parameterization of cloud formation may worsen the performance of the model with regard to precipitation (Baumberger et al., 2017). This mismatch can be caused by model over-specification, or over-tuning.

The first step in evaluating the new parameterization is to compare the outputs with a control simulation to make sure the new addition does not negatively interact with the dynamical core or spoil the contributions of rest of the parameterizations. Figure 4 shows that this is not case with the POPEM parameterization~~Figures 6C-6D and 8C-8D show that this is not case with the POPEM parameterization~~, which does not negatively affect the outputs of precipitation and temperature. Rather, both variables are now closer to the observed data than they were in the control run, especially in terms of reducing the double ITCZ, which artificially features in global models (Mechoso et al., 1995; for a recent analysis of double ITCZ in CMIP5 models see Oueslati and Bellon, 2015).~~.~~

Figure 4 about here

Figure 4A shows that there is just a slight discrepancy in the absolute difference in rainfall between the GPCP and CESM simulations (The first and the third quartiles of the distribution remain between ± 0.4 mm/day). Grid point to grid point comparison between the model and GPCP indicates the ability of CESM to reproduce the spatial distribution of precipitation. In both simulations, the CESM exhibits a good correlation coefficient (0.72 R^2) compared with the reference data (Figure 4C). The results are even better for temperature (0.88 R^2; Figure 4D).

Direct comparison of aggregated data is a standard procedure for gauging model abilities. ~~Figure 4~~Figure 5 compares two latitude-time graphs for precipitation (A) and surface temperature (B), both for the CONTROL case and for the new POPEM parameterization.

Figure 5 about here ~~Figure 4 about here~~

It is clear from Figures 5A and 6A that POPEM does alter the spatial pattern of precipitation and exerts a definite effect on the climate pattern, as the module reduces the otherwise exaggerated ITCZ precipitation in the Southern Hemisphere reported by several authors (Hwang and Frierson, 2013; Lin and Xie 2014).~~It is clear from the figure that POPEM does alter the spatial pattern of precipitation and exerts a definite effect on the climate pattern, as the module reduces the otherwise exaggerated ITCZ precipitation in the Southern Hemisphere (South East Asia and Australia). Disparities in temperature between the CONTROL and POPEM runs are apparent at high latitudes. In this case, POPEM produces lower temperatures at both poles, a result which deserves further attention.~~

~~Figure 5 about here~~

Disparities in temperature between the CONTROL and POPEM runs are apparent at high latitudes. In this case, POPEM produces lower temperatures at both poles, a result which deserves further attention (Figures 5B and 6B).

There are also important differences in precipitation in the 30N-30S band. Here POPEM reduces model bias, especially in the Southern Hemisphere and on the Tibetan Plateau. (see section 3.2 for more details). On the other hand, POPEM departs from the control simulation in the Asia-Pacific region between 10N-10S. This result reinforces the double ITCZ bias in this area.On the other hand, POPEM departs from the control simulation in the Asia-Pacific region between 10N-10S.

These results show that the POPEM parameterization generally agrees with historical data for population, and also compares well with the control simulation in the sense of addressing some of the known biases in precipitation and temperature, offering a more detailed version of CO_2 emissions at a relatively cheap computational cost. As discussed above, the CONTROL run uses global concentration values to include CO_2 on the assumption that it is well-mixed in the atmosphere (Neale et al., 2012). This assumption reduces the computational burden of the simulation but does not allow for precise emissions modeling in the future. This is an important aspect for regionalized emissions scenarios, since even if the new parameterization is not significantly better than the old approach (but no worse), it is desirable as it allows sensitivity analyses, such as evaluating the effects of the U.S. leaving the Paris agreement.

Potential applications of POPEM include not only sensitivity analyses of local CO_2 emissions policies, but also the added feature of performing tests for 'what-if' scenarios. One interesting example would be the climate response under the hypothesis that China and India –the most populated countries in the world- reach US CO_2 per capita emissions rates. Another 'what-if' scenario would be the climate response of an increasingly urbanized world. In both cases, POPEM provides a flexible framework for testing the alternative hypotheses.

The realism of the ESM will be enhanced with a fully-coupled system. Such a fully-fledged ESM will include bidirectional feedback between POPEM and CESM to evaluate the effects of climate change on population dynamics and emissions.

3.2 Validation against observational data sets

Once it has been verified that the new parameterization does not worsen the modeling, the next step in evaluating the performances is comparing the simulation outputs for both the CONTROL run and the POPEM module using actual observational data. Direct comparisons with historical data can help show whether or not a climate model correctly represents the climate of the past. However, although observational measurements are often considered the ground truth to validate models against, it is important to be aware that measurements have their own uncertainties (Tapiador et al. 2017).

Figure 7Figure 6 shows a comparison of CESM precipitation simulations for the period 1980-2000 using the GPCP. It is apparent that there is an overall consensus, even though there are differences. Despite these known biases, the model agrees with the observations on the major features of global precipitation. In Figure 6C, there is just a slight discrepancy in the absolute difference in rainfall between the GPCP and CESM simulations (Q1 and Q3 remain between ± 0.4 mm/day). Grid point to grid point comparison between the model and GPCP (a stringent comparison; Figure 6D) indicates the ability of CESM to reproduce the spatial distribution of precipitation. In both simulations, the CESM exhibits a good correlation coefficient (0.72 R^2) compared with the reference data.

Figure 7 about here Figure 6 about here

The improvements in parameterizing emissions become clearer if we focus on specific regions. For the El Niño-4 area, there are statistically-significant differences (at the 0.05 significance level) between both the CONTROL run and the POPEM modeling when compared with the reference data. This observation illustrates the limitations of the modeling and the need of advances in the parameterizations. However, for this area the correlation (R^2) between POPEM and GPCP is slightly better than CONTROL and GPCP (0.706 R^2 versus 0.692 R^2).

The real added value, however, is not in a better estimation of the totals but in the ability of POPEM to better capture the structure of the precipitation. Figure 7 shows the histograms of mean precipitation in the El Niño-4 area using the POPEM parameterization (top), the standard forcing approach (CONTROL, middle), and the reference GPCP estimates (bottom). While the CONTROL simulation severely overestimates the low end of the distribution, POPEM gives a more realistic value. This result is not apparent in the otherwise improved correlation of POPEM, and is also buried in the box plots.

El Niño-4 is important because it presents a lower variance in the SST than any other of the El Niño areas, playing a key role in identifying El Niño Modoki events (Ashok et al., 2007; Ashok and Yamagata, 2009; Yeh et al., 2009). The consequences of such events are severe disruptions in human activities due to the increased risk of droughts, heat waves, poor air quality and wildfires (McPhaden et al., 2006). Thus, precise modeling of the processes in this sector of the Pacific is extremely important.

Figure 8 about here

Another important benefit of POPEM is the reduction of the double ITCZ bias in the Southern Hemisphere. Although a small change can be inferred from Figure 7A-B, the improvement is buried in the annual mean precipitation maps. Figure 9A shows that the POPEM results are closer to observations of the intra-annual variability of precipitation, especially for the driest months (June-October).

Figure 9 about here

The figure also shows slight improvements for another two typical biases seen in CESM, namely the excess precipitation in the Tibetan Plateau (Chen and Frauenfeld, 2014; Su et al., 2013; Figure 9C) and the bias in some areas affected by the Asian-Australian monsoon (AAM), such as the Australia Top End (Meehl and Arblaster, 1998; Meehl et al. 2012; Figure 9B).

The results for the El Niño-4 area show that detailed, grid-point emissions of GHG improves the quantification of precipitation in dry areas, in agreement with our hypothesis about the benefits of locally-distributed versus global mean forcings.~~The improvements of POPEM for the El Niño-4 area show that detailed, dynamical modeling of GHG emissions is important for more precisely quantifying precipitation in dry areas, which validates the main hypothesis of the paper~~. Also, the double ITCZ example~~, this example~~ shows that the transient effects of regionalized GHG emissions may even translate into (long) 50-yr climatologies, meaning there is room for improvement in the 'rapidly mixing, well-mixed gases' forcing approach.

~~Figure 7 about here~~

Figure 8 compares the annual mean temperatures for the period 1950-2000. CESM simulations show a significant bias in high latitudes of the Northern Hemisphere (cfr. Figures 10A and 10B~~cfr. Figures 8A and 8B~~). In these areas, the model produces colder temperatures than those registered in the CRUTS reference data but this is also an issue in the CONTROL run. This deviation is also apparent in Figure 4B, where negative values lie away from the idealized regression line, and indicate further improvement of the CESM.~~This deviation is apparent in Figure 8D, where negative values lie away from the idealized regression line, and indicate further improvement of the CESM.~~

Figure 10 about here~~Figure 8 about here~~

The bias is also reproduced when compared with temperature anomalies for a specific region. Thus, for instance, CESM gives poor scores in the Barents Sea area (Figure 11; top) while POPEM obtains better results for the Bering Sea, especially in the Russian part (Figure 11; middle). Here, POPEM gives more realistic values for the period 1970-1998 but, even with the improvement, the model still overestimates the temperature anomaly.

Figure 11 about here

If we focus on global temperature anomalies, CESM simulations are able to reproduce the progressive increase in the temperature anomaly (Figure 12; top). However, the CONTROL case simulates a sharp drop at the end of the period (1990-1999), while POPEM portrays this change as smooth, in agreement with the observations.

Figure 12 about here

The differences between CONTROL and POPEM are better demonstrated when comparing land and ocean separately (Figure 12; middle and bottom). While the temperature anomalies for land are quite similar in both cases, POPEM provides a better representation of the ocean tendency from 1992 onwards, and that translates to an overall improvement (Figure 12, top).

3.3 Validation against ESPI and ONI indices

The El Niño-Southern Oscillation (ENSO) is the most dominant inter-annual climate variation in the tropics. It occurs when seasonally averaged sea surface temperature anomalies in the eastern Pacific Ocean exceed a given threshold and cause a shift in the atmospheric circulation (Trenberth 1997). Historically, the definition of ENSO does not include precipitation because of the limitations of stations (Ropelewski and Halpert, 1987), but recent work with satellites has confirmed that this phenomenon is a major driver of global precipitation variability (Haddad et al., 2004).

A major advantage of satellite-derived precipitation indices over more conventional ones is the description of the strength and position of the Walker circulation (Curtis and Adler, 2000). Under that assumption, Curtis and Adler (2000) derived three satellite-based precipitation indices: the ENSO precipitation index (ESPI); El Niño index (EI); and La Niña index (LI). Precipitation anomalies are averaged over areas of the Equatorial Pacific and Maritime Continent -where the strongest precipitation anomalies associated with ENSO are found- to construct differences or basin-wide gradients (Curtis, 2008).

Figure 13 shows a comparison of GPCP, CONTROL, and POPEM for the ESPI, EI and LI indices.

Figure 13 about here

Unfortunately, CONTROL and POPEM cases have difficulty simulating the precipitation patterns associated with ENSO. Figure 13 shows that bias increases in 82-83 and 97-98 El Niño years. The same bias emerges when comparing the EI and LI indices. In that case, the CESM model produces stronger El Niño/La Niña events than the observed data. Consequently, we can consider that CESM is unable to obtain a precise estimate of precipitation patterns, suggesting that current climate models are far from generating realistic simulations of the precipitation field (Dai, 2006).

Another widely used ENSO index is the Oceanic Niño Index (hereafter ONI). ONI was developed by the NOAA Climate Prediction Center (CPC) as the principal means for monitoring, assessing and predicting ENSO (Kousky and Higgins, 2007).

This index is defined as 3-month running-mean values of SST departures from the average in the Niño-3.4 region. It is computed from a set of homogeneous historical SST analyses (Kousky and Higgins, 2007, Smith et al. 2002).

Figure 14 about here

Figure 14 compares the ONI index for CPC, POPEM and CONTROL cases. It is clear from the figure, that POPEM produces a more realistic representation of the ENSO, especially if we focus on the 1992-1999 period. POPEM also obtains better results than CONTROL in the number of simulated el Niño events (see Table 1). The improvement is also noticeable in the intensity. The CONTROL case exhibits an overly strong ENSO -a common bias in CESM (Tang et al., 2016)- but POPEM reduces this bias (0.22^{o} C versus 0.59^{o} C).

Table 1 about here

Another important indicator is the mean duration of El Niño events. Table 1 shows that POPEM obtains better results according to observations (11 months in CPC, 10 months in POPEM, and 19 months in CONTROL).

4. Conclusions and future work

Like all models, climate models are simplified versions of the real world and therefore do not include the full complexity of the Earth system. Due to certain limitations, e.g. computational resources, or spatial and temporal resolution, climate models have to make assumptions and resort to parameterizations.

One important simplification is to use prescribed forcings instead of dynamically modeling GHG emissions. However, precise modeling of anthropogenic CO_2 emissions is important for climate change research as it allows sensitivity analyses to be performed.

Here we present a new module of gridded CO_2 emissions that is coupled with CESM. The module, denominated POPEM, computes anthropogenic CO_2 emissions by using population estimates as a proxy for disaggregating emissions beyond the national level. POPEM makes CESM use dynamical emissions data instead of fixed concentration parameters.

In terms of population and emissions, the module compares well when validated with data. Thus, POPEM's estimates for the 1950-2000 period are in general agreement with population and emission inventories from the recent past. In spite of the more realistic depiction of the actual emissions (Figure 3), issues persist. The performance of the model can be further improved in places where population projections are difficult to model. For instance, POPEM tends to underestimate emissions on the West Coast of the United States and the Anatolian Plateau, and overestimates emissions in China and Japan.

When the POPEM module is coupled with CESM to generate climatologies, the ability to successfully model precipitation and surface temperature is preserved. Moreover, the results of 50-year simulations show that the dynamical modeling of emissions produced by POPEM results in slight but noticeable differences in the resultant precipitation regime and surface temperature. Thus, dynamically modeling the emissions alters the ITCZ by reducing precipitation in the Southern Hemisphere and increasing it in the Northern Hemisphere. For particularly interesting areas, such as the El Niño-4 region, the POPEM outperforms the traditional approach.

Further work will be devoted to improving the modeling of those areas and hopefully minimizing some of the original biases of the CESM model. These include the emergence of a double ITCZ (Intertropical Convergence Zone) in CESM simulations, which is a common bias for most climate models (Oueslati and Bellon, 2015), as well as sea surface temperatures (SST) simulated by climate models, which are generally too low in the Northern Hemisphere and too high in the Southern Hemisphere (Wang et al., 2014).

Current applications of the parameterization include evaluating the effects of changes on regional policies, and a better understanding of the carbon cycle (Friedlingstein et al., 2006). Future work will be devoted to evaluating the climate response to alternative anthropogenic CO_2 emissions; to fully coupling Human-Earth subsystems; to increasing the spatial resolution of the simulations; and to refining the spatial and temporal distribution of emission estimates.

Although the version of POPEM presented here is already functional, this work is intended to be just the first step in fully coupling socioeconomic dynamics with ESMs. This will include bidirectional feedback between Human and Earth systems and the simulation of societal processes based on the internal dynamics of the model instead of using external sources to make the projections. Only within a coupled global Human-Earth system framework can we produce more realistic representations of the Earth system capturing much of the counterintuitive feedback that is missing from current models (Motesharrei et al. 2016). The success of this approach will depend on the ability of scientists from different research fields to work in an interdisciplinary framework of continuous collaboration.

~~Although the version of POPEM presented here is already functional, this work is intended to be just the first step in fully coupling socioeconomic dynamics with ESMs. Current applications of the parameterization include evaluating the effects of changes in regional policies, and a better understanding of the carbon cycle (Friedlingstein et al., 2006). Future work will be devoted to evaluating climate response to alternative anthropogenic CO_2 emissions; to increasing the spatial resolution of the simulations; and to refining the spatial and temporal distribution of emission estimates. It is envisioned that CESM~~

simulations employing an enhanced representation of societal processes will provide a more realistic depiction of the Earth System, improving the modeling of temperature, precipitation and other variables of interest.

Author Contribution

ANM and FJT contributed to experiment design, coding, analysis, manuscript writing and made the amendments suggested by the referees. RMG contributed to manuscript writing and POPEM-CESM implementation in the UCLM supercomputing center.ANM, RMG and FJT contributed to experiment design, coding, analysis, and manuscript writing.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

Funding from projects CGL2013-48367-P, CGL2016-80609-R (Ministerio de Economía y Competitividad, Ciencia e Innovación) is gratefully acknowledged. ANM acknowledges support from grant FPU 13/02798 for carrying out his PhD. We want to thank the five referees for their constructive comments and recommendations. Their comments have greatly improved the manuscript.

References

Adler, R., Sapiano, M., Huffman, G., Bolvin, D., Wang, J.-J., Gu, G., Nelkin, E., Xie, P., Chiu, L., Ferraro, R., Schneider, U. and Becker, A.: New Global Precipitation Climatology Project monthly analysis product corrects satellite data shifts, GEWEX News, 26(7–9), 2016.

Alter, R. E., Douglas, H. C., Winter, J. M. and Eltahir, E. A. B.: 20th-century regional climate change in the central United States attributed to agricultural intensification, Geophys. Res. Lett., doi:10.1002/2017GL075604, 2017.

Andres, R. J., Marland, G., Fung, I. and Matthews, E.: A 1° × 1° distribution of carbon dioxide emissions from fossil fuel consumption and cement manufacture, 1950-1990, Global Biogeochem. Cycles, 10(3), 419–429, doi:10.1029/96GB01523, 1996.

Andres, R. J., Boden, T. A. and Higdon, D. M.: Gridded uncertainty in fossil fuel carbon dioxide emission maps, a CDIAC example, Atmos. Chem. Phys., 16(23), 14979–14995, doi:10.5194/acp-16-14979-2016, 2016.

Archer, D., Eby, M., Brovkin, V., Ridgwell, A., Cao, L., Mikolajewicz, U., Caldeira, K., Matsumoto, K., Munhoven, G., Montenegro, A. and Tokos, K.: Atmospheric Lifetime of Fossil Fuel Carbon Dioxide, Annu. Rev. Earth Planet. Sci., 37(1), 117–134, doi:10.1146/annurev.earth.031208.100206, 2009.

Ashok, K. and Yamagata, T.: Climate change: The El Nĩo with a difference, Nature, 461(7263), 481–484, doi:10.1038/461481a, 2009.

Ashok, K., Behera, S. K., Rao, S. A., Weng, H. and Yamagata, T.: El Niño Modoki and its possible teleconnection, J. Geophys. Res. Ocean., 112(11), C11007, doi:10.1029/2006JC003798, 2007.

Bacmeister, J. T., Wehner, M. F., Neale, R. B., Gettelman, A., Hannay, C., Lauritzen, P. H., Caron, J. M. and Truesdale, J. E.: Exploratory high-resolution climate simulations using the community atmosphere model (CAM), J. Clim., 27(9), 3073–3099, doi:10.1175/JCLI-D-13-00387.1, 2014.

Baker, N. C., Taylor, P. C., Baker, N. C. and Taylor, P. C.: A Framework for Evaluating Climate Model Performance Metrics, J. Clim., 29(5), 1773–1782, doi:10.1175/JCLI-D-15-0114.1, 2016.

Barnett, T. P., Pierce, D. W., Hidalgo, H. G., Bonfils, C., Santer, B. D., Das, T., Bala, G., Wood, A. W., Nozawa, T., Mirin, A. A., Cayan, D. R. and Dettinger, M. D.: Human-induced changes in the hydrology of the Western United States, Science (80-.)., 319(5866), 1080–1083, doi:10.1126/science.1152538, 2008.

Baumberger, C., Knutti, R. and Hirsch Hadorn, G.: Building confidence in climate model projections: an analysis of inferences from fit, Wiley Interdiscip. Rev. Clim. Chang., 8(3), e454, doi:10.1002/wcc.454, 2017.

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(3), 687–720, doi:10.5194/gmd-6-687-2013, 2013.

Boden, T. A., Marland, G. and Andres, R.: Global, Regional, and National Fossil-Fuel CO2 Emissions, Carbon Dioxide Inf. Anal. Center, Oak Ridge Natl. Lab. U.S. Dep. Energy, Oak Ridge, Tenn., USA, doi:https://doi.org/10.3334/CDIAC/00001_V2017, 2017, 2017.

Boville, B. A. and Gent, P. R.: The NCAR climate system model, version one, J. Clim., 11(6), 1115–1130, doi:10.1175/1520-0442(1998)011<1115:TNCSMV>2.0.CO;2, 1998.

Brown, P. T., Li, W., Cordero, E. C. and Mauget, S. A.: Comparing the model-simulated global warming signal to observations using empirical estimates of unforced noise, Sci. Rep., 5(1), 9957, doi:10.1038/srep09957, 2015.

Chen, L. and Frauenfeld, O. W.: A comprehensive evaluation of precipitation simulations over China based on CMIP5 multimodel ensemble projections, J. Geophys. Res., 119(10), 5767–5786, doi:10.1002/2013JD021190, 2014.

Christensen, J. H. and Boberg, F.: Temperature dependent climate projection deficiencies in CMIP5 models, Geophys. Res. Lett., 39(24), doi:10.1029/2012GL053650, 2012.

Collins, W. D., Bitz, C. M., Blackmon, M. L., Bonan, G. B., Bretherton, C. S., Carton, J. A., Chang, P., Doney, S. C., Hack, J. J., Henderson, T. B., Kiehl, J. T., Large, W. G., McKenna, D. S., Santer, B. D. and Smith, R. D.: The Community Climate System Model version 3 (CCSM3), J. Clim., 19(11), 2122–2143, doi:10.1175/JCLI3761.1, 2006.

Craig, A. P., Vertenstein, M. and Jacob, R.: A new flexible coupler for earth system modeling developed for CCSM4 and CESM1, Int. J. High Perform. Comput. Appl., 26(1), 31–42, doi:10.1177/1094342011428141, 2012.

Crutzen, P. J.: Geology of mankind, Nature, 415(6867), 23, doi:10.1038/415023a, 2002.

Edwards, P. N.: History of climate modeling, Wiley Interdiscip. Rev. Clim. Chang., 2(1), 128–139, doi:10.1002/wcc.95, 2011.

Curtis, S.: The El Niño–Southern Oscillation and Global Precipitation, Geogr. Compass, 2(3), 600–619, doi:10.1111/j.1749-8198.2008.00105.x, 2008.

Curtis, S., Adler, R., Curtis, S. and Adler, R.: ENSO Indices Based on Patterns of Satellite-Derived Precipitation, J. Clim., 13(15), 2786–2793, doi:10.1175/1520-0442(2000)013<2786:EIBOPO>2.0.CO;2, 2000.

Dai, A.: Precipitation characteristics in eighteen coupled climate models, J. Clim., 19(18), 4605–4630, doi:10.1175/JCLI3884.1, 2006.

Davey, M., Huddleston, M., Sperber, K., Braconnot, P., Bryan, F., Chen, D., Colman, R., Cooper, C., Cubasch, U., Delecluse, P., DeWitt, D., Fairhead, L., Flato, G., Gordon, C., Hogan, T., Ji, M., Kimoto, M., Kitoh, A., Knutson, T., Latif, M., Le Treut, H., Li, T., Manabe, S., Mechoso, C., Meehl, G., Power, S., Roeckner, E., Terray, L., Vintzileos, A., Voss, R., Wang, B., Washington, W., Yoshikawa, I., Yu, J., Yukimoto, S. and Zebiak, S.: STOIC: A study of coupled model climatology and variability in tropical ocean regions, Clim. Dyn., 18(5), 403–420, doi:10.1007/s00382-001-0188-6, 2002.

Edwards, P. N.: History of climate modeling, Wiley Interdiscip. Rev. Clim. Chang., 2(1), 128–139, doi:10.1002/wcc.95, 2011.

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(5), 1937–1958, doi:10.5194/gmd-9-1937-2016, 2016.

Flato, G. M.: Earth system models: an overview, Wiley Interdiscip. Rev. Clim. Chang., 2(6), 783–800, doi:10.1002/wcc.148, 2011.

Fogli, P. G. and Iovino, D.: CMCC–CESM–NEMO: Toward the new CMCC Earth System Model, C. Res. Rep. RP0248, 19 [online] Available from: http://www.cmcc.it/wp-content/uploads/2015/02/rp0248-ans-12-2014.pdf, 2014.

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., von Bloh, W., Brovkin, V., Cadule, P., Doney, S., Eby, M., Fung, I., Bala, G., John, J., Jones, C., Joos, F., Kato, T., Kawamiya, M., Knorr, W., Lindsay, K., Matthews, H. D., Raddatz, T., Rayner, P., Reick, C., Roeckner, E., Schnitzler, K. G., Schnur, R., Strassmann, K., Weaver, A. J., Yoshikawa, C. and Zeng, N.: Climate-carbon cycle feedback analysis: Results from the C4MIP model intercomparison, J. Clim., 19(14), 3337–3353, doi:10.1175/JCLI3800.1, 2006.

Gately, C. K., Hutyra, L. R., Wing, I. S. and Brondfield, M. N.: A bottom up approach to on-road CO2 emissions estimates: Improved spatial accuracy and applications for regional planning, Environ. Sci. Technol., 47(5), 2423–2430, doi:10.1021/es304238v, 2013.

Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., Lawrence, D. M., Neale, R. B., Rasch, P. J., Vertenstein, M., Worley, P. H., Yang, Z. L. and Zhang, M.: The community climate system model version 4, J. Clim., 24(19), 4973–4991, doi:10.1175/2011JCLI4083.1, 2011.

Giorgetta, M. A., Jungclaus, J., Reick, C. H., Legutke, S., Bader, J., Böttinger, M., Brovkin, V., Crueger, T., Esch, M., Fieg, K., Glushak, K., Gayler, V., Haak, H., Hollweg, H.-D., Ilyina, T., Kinne, S., Kornblueh, L., Matei, D., Mauritsen, T., Mikolajewicz, U., Mueller, W., Notz, D., Pithan, F., Raddatz, T., Rast, S., Redler, R., Roeckner, E., Schmidt, H., Schnur, R., Segschneider, J., Six, K. D., Stockhause, M., Timmreck, C., Wegner, J., Widmann, H., Wieners, K.-H., Claussen, M., Marotzke, J. and Stevens, B.: Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5, J. Adv. Model. Earth Syst., 5(3), 572–597, doi:10.1002/jame.20038, 2013.

Grandey, B. S., Cheng, H. and Wang, C.: Transient climate impacts for scenarios of aerosol emissions from Asia: A story of coal versus gas, J. Clim., 29(8), 2849–2867, doi:10.1175/JCLI-D-15-0555.1, 2016.

Guo, L., Highwood, E. J., Shaffrey, L. C. and Turner, A. G.: The effect of regional changes in anthropogenic aerosols on rainfall of the East Asian Summer Monsoon, Atmos. Chem. Phys., 13(3), 1521–1534, doi:10.5194/acp-13-1521-2013, 2013.

Gurney, K. R., Mendoza, D. L., Zhou, Y., Fischer, M. L., Miller, C. C., Geethakumar, S. and de la Rue du Can, S.: High Resolution Fossil Fuel Combustion CO 2 Emission Fluxes for the United States, Environ. Sci. Technol., 43(14), 5535–5541, doi:10.1021/es900806c, 2009.

Haddad, Z. S., Meagher, J. P., Adler, R. F., Smith, E. A., Im, E. and Durden, S. L.: Global variability of precipitation according to the Tropical Rainfall Measuring Mission, J. Geophys. Res. D Atmos., 109(17), D17103, doi:10.1029/2004JD004607, 2004.

Hansen, J. and Lebedeff, S.: Global trends of measured surface air temperature, J. Geophys. Res., 92(D11), 13345–13372, doi:10.1029/JD092iD11p13345, 1987.

Hansen, J., Ruedy, R., Sato, M. and Reynolds, R.: Global surface air temperature in 1995: Return to pre-Pinatubo level, Geophys. Res. Lett., 23(13), 1665–1668, doi:10.1029/96GL01040, 1996.

Hansen, J., Ruedy, R., Sato, M. and Lo, K.: Global surface temperature change, Rev. Geophys., 48(4), RG4004, doi:10.1029/2010RG000345, 2010.

Hao, Z., AghaKouchak, A. and Phillips, T. J.: Changes in concurrent monthly precipitation and temperature extremes, Environ. Res. Lett., 8(3), 34014, doi:10.1088/1748-9326/8/3/034014, 2013.

Hargreaves, J. C.: Skill and uncertainty in climate models, Wiley Interdiscip. Rev. Clim. Chang., 1(4), 556–564, doi:10.1002/wcc.58, 2010.

Hargreaves, J. C. and Annan, J. D.: Can we trust climate models?, Wiley Interdiscip. Rev. Clim. Chang., 5(4), 435–440, doi:10.1002/wcc.288, 2014.

Harris, I., Jones, P. D., Osborn, T. J. and Lister, D. H.: Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset, Int. J. Climatol., 34(3), 623–642, doi:10.1002/joc.3711, 2014.

Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., Oki, R., Nakamura, K. and Iguchi, T.: The global precipitation measurement mission, Bull. Am. Meteorol. Soc., 95(5), 701–722, doi:10.1175/BAMS-D-13-00164.1, 2014.

Hourdin, F., Mauritsen, T., Gettelman, A., Golaz, J.-C., Balaji, V., Duan, Q., Folini, D., Ji, D., Klocke, D., Qian, Y., Rauser, F., Rio, C., Tomassini, L., Watanabe, M. and Williamson, D.: The Art and Science of Climate Model Tuning, Bull. Am. Meteorol. Soc., 98(3), 589–602, doi:10.1175/BAMS-D-15-00135.1, 2017.

Huang, X. and Ullrich, P. A.: Irrigation impacts on California's climate with the variable-resolution CESM, J. Adv. Model. Earth Syst., 8(3), 1151–1163, doi:10.1002/2016MS000656, 2016.

Hunke, E. C. and Lipscomb, W. H.: CICE: the Los Alamos sea ice model user's manual, version 4. Los Alamos National Laboratory Tech. Rep. [online] Available from: https://www.osti.gov/scitech/biblio/1364126, 2008.

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., Lamarque, J.-F. F., Large, W. G., Lawrence, D., Lindsay, K., Lipscomb, W. H., Long, M. C., Mahowald, N., Marsh, D. R., Neale, R. B., Rasch, P., Vavrus, S., Vertenstein, M., Bader, D., Collins, W. D., Hack, J. J., Kiehl, J. and Marshall, S.: The Community Earth System Model: A Framework for Collaborative Research, Bull. Am. Meteorol. Soc., 94(9), 1339–1360, doi:10.1175/BAMS-D-12-00121.1, 2013.

Hwang, Y.-T. and Frierson, D. M. W.: Link between the double-Intertropical Convergence Zone problem and cloud biases over the Southern Ocean., Proc. Natl. Acad. Sci. U. S. A., 110(13), 4935–40, doi:10.1073/pnas.1213302110, 2013.

IPCC: Climate Change 2013 - The Physical Science Basis, edited by Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge., 2014a.

IPCC: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by R. K. Pachauri and L. A. Meyer, Geneva, Switzerland, 151 pp., 2014b.

Justin Small, R., Curchitser, E., Hedstrom, K., Kauffman, B. and Large, W. G.: The Benguela upwelling system: Quantifying the sensitivity to resolution and coastal wind representation in a global climate model, J. Clim., 28(23), 9409–9432, doi:10.1175/JCLI-D-15-0192.1, 2015.

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. F., 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. Am. Meteorol. Soc., 96(8), 1333–1349, doi:10.1175/BAMS-D-13-00255.1, 2015.

Kerbyson, D. J. and Jones, P. W.: A Performance Model of the Parallel Ocean Program, Int. J. High Perform. Comput. Appl., 19(3), 261–276, doi:10.1177/1094342005056114, 2005.

Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G. and Kirschbaum, D. B.: So, how much of the Earth's surface is covered by rain gauges?, Bull. Am. Meteorol. Soc., 98(1), 69–78, doi:10.1175/BAMS-D-14-00283.1, 2017.

Kousky, V. E. and Higgins, R. W.: An Alert Classification System for Monitoring and Assessing the ENSO Cycle, Weather Forecast., 22(2), 353–371, doi:10.1175/WAF987.1, 2007.

Ladyman, J., Lambert, J. and Wiesner, K.: What is a complex system?, Eur. J. Philos. Sci., 3(1), 33–67, doi:10.1007/s13194-012-0056-8, 2013.

Lahsen, M.: Seductive Simulations? Uncertainty Distribution Around Climate Models, Soc. Stud. Sci., 35(6), 895–922, doi:10.1177/0306312705053049, 2005.

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., Zeng, X., Yang, Z.-L., Levis, S., Sakaguchi, K., Bonan, G. B. and Slater, A. G.: Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model, J. Adv. Model. Earth Syst., 3 [online] Available from: http://doi.wiley.com/10.1029/2011MS00045, 2011.

Lehner, F., Joos, F., Raible, C. C., Mignot, J., Born, A., Keller, K. M. and Stocker, T. F.: Climate and carbon cycle dynamics in a CESM simulation from 850 to 2100 CE, Earth Syst. Dyn., 6(2), 411–434, doi:10.5194/esd-6-411-2015, 2015.

Levis, S., Gordon, B. B., Erik Kluzek, Thornton, P. E., Jones, A., Sacks, W. J. and Kucharik, C. J.: Interactive Crop Management in the Community Earth System Model (CESM1): Seasonal influences on land-atmosphere fluxes, J. Clim., 25(14), 4839–4859, doi:10.1175/JCLI-D-11-00446.1, 2012.

Li, G. and Xie, S. P.: Tropical biases in CMIP5 multimodel ensemble: The excessive equatorial pacific cold tongue and double ITCZ problems, J. Clim., 27(4), 1765–1780, doi:10.1175/JCLI-D-13-00337.1, 2014.

Lindsay, K., Bonan, G. B., Doney, S. C., Hoffman, F. M., Lawrence, D. M., Long, M. C., Mahowald, N. M., Keith Moore, J., Randerson, J. T., Thornton, P. E., Lindsay, K., Bonan, G. B., Doney, S. C., Hoffman, F. M., Lawrence, D. M., Long, M. C., Mahowald, N. M., Moore, J. K., Randerson, J. T. and Thornton, P. E.: Preindustrial-Control and Twentieth-Century Carbon Cycle Experiments with the Earth System Model CESM1(BGC), J. Clim., 27(24), 8981–9005, doi:10.1175/JCLI-D-12-00565.1, 2014.

Lipscomb, W. H., Fyke, J. G., Vizcaíno, M., Sacks, W. J., Wolfe, J., Vertenstein, M., Craig, A., Kluzek, E. and Lawrence, D. M.: Implementation and initial evaluation of the glimmer community ice sheet model in the community earth system model, J. Clim., 26(19), 7352–7371, doi:10.1175/JCLI-D-12-00557.1, 2013.

Liu, X., Easter, R. C., Ghan, S. J., Zaveri, R., Rasch, P., Shi, X., Lamarque, J. F., Gettelman, A., Morrison, H., Vitt, F., Conley, A., Park, S., Neale, R., Hannay, C., Ekman, A. M. L., Hess, P., Mahowald, N., Collins, W., Iacono, M. J., Bretherton, C. S., Flanner, M. G. and Mitchell, D.: Toward a minimal representation of aerosols in climate models: Description and evaluation in the Community Atmosphere Model CAM5, Geosci. Model Dev., 5(3), 709–739, doi:10.5194/gmd-5-709-2012, 2012.

Liu, Z., Mehran, A., Phillips, T. J. and AghaKouchak, A.: Seasonal and regional biases in CMIP5 precipitation simulations, Clim. Res., 60(1), 35–50, doi:10.3354/cr01221, 2014.

Lorenz, E. N.: Deterministic Nonperiodic Flow, J. Atmos. Sci., 20(2), 130–141, doi:10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2, 1963.

Martin, G. M., Bellouin, N., Collins, W. J., Culverwell, I. D., Halloran, P. R., Hardiman, S. C., Hinton, T. J., Jones, C. D., McDonald, R. E., McLaren, A. J., O'Connor, F. M., Roberts, M. J., Rodriguez, J. M., Woodward, S., Best, M. J., Brooks, M. E., Brown, A. R., Butchart, N., Dearden, C., Derbyshire, S. H., Dharssi, I., Doutriaux-Boucher, M., Edwards, J. M., Falloon, P. D., Gedney, N., Gray, L. J., Hewitt, H. T., Hobson, M., Huddleston, M. R., Hughes, J., Ineson, S., Ingram, W. J., James, P. M., Johns, T. C., Johnson, C. E., Jones, A., Jones, C. P., Joshi, M. M., Keen, A. B., Liddicoat, S., Lock, A. P., Maidens, A. V., Manners, J. C., Milton, S. F., Rae, J. G. L., Ridley, J. K., Sellar, A., Senior, C. A., Totterdell, I. J., Verhoef, A., Vidale, P. L. and Wiltshire, A.: The HadGEM2 family of Met Office Unified Model climate configurations, Geosci. Model Dev., 4(3), 723–757, doi:10.5194/gmd-4-723-2011, 2011.

Mauritsen, T., Stevens, B., Roeckner, E., Crueger, T., Esch, M., Giorgetta, M., Haak, H., Jungclaus, J., Klocke, D., Matei, D., Mikolajewicz, U., Notz, D., Pincus, R., Schmidt, H. and Tomassini, L.: Tuning the climate of a global model, J. Adv. Model. Earth Syst., 4(8), n/a-n/a, doi:10.1029/2012MS000154, 2012.

McPhaden, M. J., Zebiak, S. E. and Glantz, M. H.: ENSO as an Integrating Concept in Earth Science, Science (80-.)., 314(5806), 1740–1745, doi:10.1126/science.1132588, 2006.

Mechoso, C. R., Robertson, A. W., Barth, N., Davey, M. K., Delecluse, P., Gent, P. R., Ineson, S., Kirtman, B., Latif, M., Treut, H. Le, Nagai, T., Neelin, J. D., Philander, S. G. H., Polcher, J., Schopf, P. S., Stockdale, T., Suarez, M. J., Terray, L., Thual, O. and Tribbia, J. J.: The Seasonal Cycle over the Tropical Pacific in Coupled Ocean–Atmosphere General Circulation Models, Mon. Weather Rev., 123(9), 2825–2838, doi:10.1175/1520-0493(1995)123<2825:TSCOTT>2.0.CO;2, 1995.

Meehl, G. A. and Arblaster, J. M.: The Asian-Australian monsoon and El Nino-Southern Oscillation in the NCAR climate system model, J. Clim., 11(6), 1356–1385, doi:10.1175/1520-0442(1998)011<1356:TAAMAE>2.0.CO;2, 1998.

Meehl, G. A., Arblaster, J. M., Caron, J. M., Annamalai, H., Jochum, M., Chakraborty, A. and Murtugudde, R.: Monsoon regimes and processes in CCSM4. Part I: The Asian-Australian monsoon, J. Clim., 25(8), 2583–2608, doi:10.1175/JCLI-D-11-00184.1, 2012.

Monier, E., Scott, J. R., Sokolov, A. P., Forest, C. E. and Schlosser, C. A.: An integrated assessment modeling framework for uncertainty studies in global and regional climate change: The MIT IGSM-CAM (version 1.0), Geosci. Model Dev., 6(6), 2063–2085, doi:10.5194/gmd-6-2063-2013, 2013.

Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., van Vuuren, D. P., Carter, T. R., Emori, S., Kainuma, M., Kram, T., Meehl, G. A., Mitchell, J. F. B., Nakicenovic, N., Riahi, K., Smith, S. J., Stouffer, R. J., Thomson, A. M., Weyant, J. P. and Wilbanks, T. J.: The next generation of scenarios for climate change research and assessment., Nature, 463(7282), 747–56, doi:10.1038/nature08823, 2010.

Motesharrei, S., Rivas, J. and Kalnay, E.: Human and nature dynamics (HANDY): Modeling inequality and use of resources in the collapse or sustainability of societies, Ecol. Econ., 101, 90–102, doi:10.1016/j.ecolecon.2014.02.014, 2014.

Motesharrei, S., Rivas, J., Kalnay, E., Asrar, G. R., Busalacchi, A. J., Cahalan, R. F., Cane, M. A., Colwell, R. R., Feng, K., Franklin, R. S., Hubacek, K., Miralles-Wilhelm, F., Miyoshi, T., Ruth, M., Sagdeev, R., Shirmohammadi, A., Shukla, J., Srebric, J., Yakovenko, V. M. and Zeng, N.: Modeling Sustainability: Population, Inequality, Consumption, and Bidirectional Coupling of the Earth and Human Systems, Natl. Sci. Rev., 3(4), nww081, doi:10.1093/nsr/nww081, 2016.

Müller-Hansen, F., Schlüter, M., Mäs, M., Hegselmann, R., Donges, J. F., Kolb, J. J., Thonicke, K. and Heitzig, J.: How to represent human behavior and decision making in Earth system models? A guide to techniques and approaches, Earth Syst. Dyn. Discuss., 1–53, doi:10.5194/esd-2017-18, 2017.

Nasrollahi, N., Aghakouchak, A., Cheng, L., Damberg, L., Phillips, T. J., Miao, C., Hsu, K. and Sorooshian, S.: How well do CMIP5 climate simulations replicate historical trends and patterns of meteorological droughts?, Water Resour. Res., 51(4), 2847–2864, doi:10.1002/2014WR016318, 2015.

Navarro, A., Moreno, R., Jiménez-Alcázar, A. and Tapiador, F. J.: Coupling population dynamics with earth system models: the POPEM model, Environ. Sci. Pollut. Res., in press, doi:10.1007/s11356-017-0127-7, 2017.

Neale, R. B., Chen, C., Lauritzen, P. H., Williamson, D. L., Conley, A. J., Smith, A. K., Mills, M. and Morrison, H.: Description of the NCAR Community Atmosphere Model (CAM 5 . 0). [online] Available from: https://www.ccsm.ucar.edu/models/ccsm4.0/cam/docs/description/cam4_desc.pdf, 2012.

Neely, R. R., Conley, A. J., Vitt, F. and Lamarque, J. F.: A consistent prescription of stratospheric aerosol for both radiation and chemistry in the Community Earth System Model (CESM1), Geosci. Model Dev., 9(7), 2459–2470, doi:10.5194/gmd-9-2459-2016, 2016.

Nevison, C., Hess, P., Riddick, S. and Ward, D.: Denitrification, leaching, and river nitrogen export in the Community Earth System Model, J. Adv. Model. Earth Syst., 8(1), 272–291, doi:10.1002/2015MS000573, 2016.

New, M., Hulme, M. and Jones, P.: Representing twentieth-century space-time climate variability. Part I: Development of a 1961-90 mean monthly terrestrial climatology, J. Clim., 12(2–3), 829–856, doi:10.1175/1520-0442(1999)012<0829:RTCSTC>2.0.CO;2, 1999.

New, M., Hulme, M., Jones, P., New, M., Hulme, M. and Jones, P.: Representing Twentieth-Century Space–Time Climate Variability. Part II: Development of 1901–96 Monthly Grids of Terrestrial Surface Climate, J. Clim., 13(13), 2217–2238, doi:10.1175/1520-0442(2000)013<2217:RTCSTC>2.0.CO;2, 2000.

Nobre, C., Brasseur, G. P., Shapiro, M. A., Lahsen, M., Brunet, G., Busalacchi, A. J., Hibbard, K., Seitzinger, S., Noone, K., Ometto, J. P., Nobre, C., Brasseur, G. P., Shapiro, M. A., Lahsen, M., Brunet, G., Busalacchi, A. J., Hibbard, K., Seitzinger, S., Noone, K. and Ometto, J. P.: Addressing the Complexity of the Earth System, Bull. Am. Meteorol. Soc., 91(10), 1389–1396, doi:10.1175/2010BAMS3012.1, 2010.

Oda, T., Maksyutov, S. and Andres, R. J.: The Open-source Data Inventory for Anthropogenic CO2 , version 2016 (ODIAC2016): a global monthly fossil fuel CO2 gridded emissions data product for tracer transport simulations and surface flux inversions, Earth Syst. Sci. Data, 10(1), 87–107, doi:10.5194/essd-10-87-2018, 2018.

Oleson, K. W., Bonan, G. B., Feddema, J. and Jackson, T.: An examination of urban heat island characteristics in a global climate model, Int. J. Climatol., 31(12), 1848–1865, doi:10.1002/joc.2201, 2011.

Oueslati, B. and Bellon, G.: The double ITCZ bias in CMIP5 models: interaction between SST, large-scale circulation and precipitation, Clim. Dyn., 44(3–4), 585–607, doi:10.1007/s00382-015-2468-6, 2015.

Peng-Fei, L., Xiao-Li, F. and Juan-Juan, L.: Historical Trends in Surface Air Temperature Estimated by Ensemble Empirical Mode Decomposition and Least Squares Linear Fitting, Atmos. Ocean. Sci. Lett., 8(1), 10–16, doi:10.3878/AOSL20140064, 2015.

Pincus, R., Batstone, C. P., Patrick Hofmann, R. J., Taylor, K. E. and Glecker, P. J.: Evaluating the present-day simulation of clouds, precipitation, and radiation in climate models, J. Geophys. Res. Atmos., 113(14), D14209, doi:10.1029/2007JD009334, 2008.

Prather, M. J.: Lifetimes and time scales in atmospheric chemistry, Philos. Trans. R. Soc. A Math. Phys. Eng. Sci., 365(1856), 1705–1726, doi:10.1098/rsta.2007.2040, 2007.

Rayner, P. J., Raupach, M. R., Paget, M., Peylin, P. and Koffi, E.: A new global gridded data set of CO 2 emissions from fossil fuel combustion: Methodology and evaluation, J. Geophys. Res., 115(D19), D19306, doi:10.1029/2009JD013439, 2010.

Richter, I.: Climate model biases in the eastern tropical oceans: Causes, impacts and ways forward, Wiley Interdiscip. Rev. Clim. Chang., 6(3), 345–358, doi:10.1002/wcc.338, 2015.

Rind, D.: Complexity and Climate, Science (80-.)., 284(5411), 105–107, doi:10.1126/science.284.5411.105, 1999.

Robinson, D. T., Di Vittorio, A., Alexander, P., Arneth, A., Barton, C. M., Brown, D. G., Kettner, A., Lemmen, C., O'Neill, B. C., Janssen, M., Pugh, T. A. M., Rabin, S. S., Rounsevell, M., Syvitski, J. P., Ullah, I. and Verburg, P. H.: Modelling feedbacks between human and natural processes in the land system, Earth Syst. Dyn. Discuss., 1–47, doi:10.5194/esd-2017-68, 2017.

Ropelewski, C. F. and Halpert, M. S.: Global and Regional Scale Precipitation Patterns Associated with the El Niño/Southern Oscillation, Mon. Weather Rev., 115(8), 1606–1626, doi:10.1175/1520-0493(1987)115<1606:GARSPP>2.0.CO;2, 1987.

Ruth, M., Kalnay, E., Zeng, N., Franklin, R. S., Rivas, J. and Miralles-Wilhelm, F.: Sustainable prosperity and societal transitions: Long-term modeling for anticipatory management, Environ. Innov. Soc. Transitions, 1(1), 160–165, doi:10.1016/j.eist.2011.03.004, 2011.

Sanderson, B. M., Xu, Y., Tebaldi, C., Wehner, M., O'neill, B., Jahn, A., Pendergrass, A. G., Lehner, F., Strand, W. G., Lin, L., Knutti, R. and Francois Lamarque, J.: Community climate simulations to assess avoided impacts in 1.5 and 2°C futures, Earth Syst. Dyn., 8(3), 827–847, doi:10.5194/esd-8-827-2017, 2017.

Sarofim, M. C. and Reilly, J. M.: Applications of integrated assessment modeling to climate change, Wiley Interdiscip. Rev. Clim. Chang., 2(1), 27–44, doi:10.1002/wcc.93, 2011.

Schellnhuber, H. J.: Earth System Science and the second Copernican revolution, Nature, 402, C19–C23, doi:10.1038/35011515, 1999.

Schmidt, G. A., Kelley, M., Nazarenko, L., Ruedy, R., Russell, G. L., Aleinov, I., Bauer, M., Bauer, S. E., Bhat, M. K., Bleck, R., Canuto, V., Chen, Y. H., Cheng, Y., Clune, T. L., Del Genio, A., De Fainchtein, R., Faluvegi, G., Hansen, J. E., Healy, R. J., Kiang, N. Y., Koch, D., Lacis, A. A., Legrande, A. N., Lerner, J., Lo, K. K., Matthews, E. E., Menon, S., Miller, R. L., Oinas, V., Oloso, A. O., Perlwitz, J. P., Puma, M. J., Putman, W. M., Rind, D., Romanou, A., Sato, M., Shindell, D. T., Sun, S., Syed, R. A., Tausnev, N., Tsigaridis, K., Unger, N., Voulgarakis, A., Yao, M. S. and Zhang, J.: Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive, J. Adv. Model. Earth Syst., 6(1), 141–184, doi:10.1002/2013MS000265, 2014.

Shukla, J., Palmer, T. N., Hagedorn, R., Hoskins, B., Kinter, J., Marotzke, J., Miller, M., Slingo, J., Shukla, J., Palmer, T. N., Hagedorn, R., Hoskins, B., Kinter, J., Marotzke, J., Miller, M. and Slingo, J.: Toward a New Generation of World Climate Research and Computing Facilities, Bull. Am. Meteorol. Soc., 91(10), 1407–1412, doi:10.1175/2010BAMS2900.1, 2010.

Smith, M. J., Palmer, P. I., Purves, D. W., Vanderwel, M. C., Lyutsarev, V., Calderhead, B., Joppa, L. N., Bishop, C. M. and Emmott, S.: Changing how earth system modeling is done to provide more useful information for decision making, science, and society, Bull. Am. Meteorol. Soc., 95(9), 1453–1464, doi:10.1175/BAMS-D-13-00080.1, 2014.

Stanfield, R. E., Jiang, J. H., Dong, X., Xi, B., Su, H., Donner, L., Rotstayn, L., Wu, T., Cole, J. and Shindo, E.: A quantitative assessment of precipitation associated with the ITCZ in the CMIP5 GCM simulations, Clim. Dyn., 47(5–6), 1863–1880, doi:10.1007/s00382-015-2937-y, 2016.

Smith, T. M., Reynolds, R. W., Smith, T. M. and Reynolds, R. W.: Extended Reconstruction of Global Sea Surface Temperatures Based on COADS Data (1854–1997), J. Clim., 16(10), 1495–1510, doi:10.1175/1520-0442-16.10.1495, 2003.

Stanfield, R. E., Jiang, J. H., Dong, X., Xi, B., Su, H., Donner, L., Rotstayn, L., Wu, T., Cole, J. and Shindo, E.: A quantitative assessment of precipitation associated with the ITCZ in the CMIP5 GCM simulations, Clim. Dyn., 47(5–6), 1863–1880, doi:10.1007/s00382-015-2937-y, 2016.

Steven, B. and Bony, S.: What Are Climate Models Missing?, Science (80-.)., 340(6136), 1053–1054 [online] Available from: http://science.sciencemag.org/content/340/6136/1053 (Accessed 18 June 2017), 2013.

Su, F., Duan, X., Chen, D., Hao, Z. and Cuo, L.: Evaluation of the global climate models in the CMIP5 over the Tibetan Plateau, J. Clim., 26(10), 3187–3208, doi:10.1175/JCLI-D-12-00321.1, 2013.

Tang, Y., Li, L., Dong, W. and Wang, B.: Tracing the source of ENSO simulation differences to the atmospheric component of two CGCMs, Atmos. Sci. Lett., 17(2), 155–161, doi:10.1002/asl.637, 2016.

Tapiador, F. J.: A joint estimate of the precipitation climate signal in Europe using eight regional models and five observational datasets, J. Clim., 23(7), 1719–1738, doi:10.1175/2009JCLI2956.1, 2010.

Tapiador, F. J., Turk, F. J., Petersen, W., Hou, A. Y., García-Ortega, E., Machado, L. A. T., Angelis, C. F., Salio, P., Kidd, C., Huffman, G. J. and de Castro, M.: Global precipitation measurement: Methods, datasets and applications, Atmos. Res., 104–105, 70–97, doi:10.1016/j.atmosres.2011.10.021, 2012.

Tapiador, F. J., Navarro, A., Levizzani, V., García-Ortega, E., Huffman, G. J., Kidd, C., Kucera, P. A., Kummerow, C. D., Masunaga, H., Petersen, W. A., Roca, R., Sánchez, J.-L., Tao, W.-K. and Turk, F. J.: Global precipitation measurements for validating climate models, Atmos. Res., doi:10.1016/j.atmosres.2017.06.021, 2017.

Tapiador, F. J., Navarro, A., Jiménez, A., Moreno, R. and García-Ortega, E.: Discrepancies with Satellite Observations in the Spatial Structure of Global Precipitation as Derived from Global Climate Models, Q. J. R. Meteorol. Soc., doi:10.1002/qj.3289, 2018.

Taylor, K. E., Stouffer, R. J., Meehl, G. A., Taylor, K. E., Stouffer, R. J. and Meehl, G. A.: An Overview of CMIP5 and the Experiment Design, Bull. Am. Meteorol. Soc., 93(4), 485–498, doi:10.1175/BAMS-D-11-00094.1, 2012.

Tilmes, S., Lamarque, J. F., Emmons, L. K., Kinnison, D. E., Ma, P. L., Liu, X., Ghan, S., Bardeen, C., Arnold, S., Deeter, M., Vitt, F., Ryerson, T., Elkins, J. W., Moore, F., Spackman, J. R. and Val Martin, M.: Description and evaluation of tropospheric chemistry and aerosols in the Community Earth System Model (CESM1.2), Geosci. Model Dev., 8(5), 1395–1426, doi:10.5194/gmd-8-1395-2015, 2015.

Trenberth, K. E.: The Definition of El Niño, Bull. Am. Meteorol. Soc., 78(12), 2771–2777, doi:10.1175/1520-0477(1997)078<2771:TDOENO>2.0.CO;2, 1997.

Trenberth, K. E., Zhang, Y. and Fasullo, J. T.: Relationships among top-of-atmosphere radiation and atmospheric state variables in observations and CESM, J. Geophys. Res., 120(19), 10074–10090, doi:10.1002/2015JD023381, 2015.

Trenberth, K. E., Zhang, Y. and Gehne, M.: Intermittency in Precipitation: Duration, Frequency, Intensity, and Amounts Using Hourly Data, J. Hydrometeorol., 18(5), 1393–1412, doi:10.1175/JHM-D-16-0263.1, 2017.

Val Martin, M., Heald, C. L. and Arnold, S. R.: Coupling dry deposition to vegetation phenology in the Community Earth System Model: Implications for the simulation of surface O3, Geophys. Res. Lett., 41(8), 2988–2996, doi:10.1002/2014GL059651, 2014.

van Vuuren, D. P. and Carter, T. R.: Climate and socio-economic scenarios for climate change research and assessment: reconciling the new with the old, Clim. Change, 122(3), 415–429, doi:10.1007/s10584-013-0974-2, 2014.

van Vuuren, D. P., Lowe, J., Stehfest, E., Gohar, L., Hof, A. F., Hope, C., Warren, R., Meinshausen, M. and Plattner, G.-K.: How well do integrated assessment models simulate climate change?, Clim. Change, 104(2), 255–285, doi:10.1007/s10584-009-9764-2, 2011.

van Vuuren, D. P., Bayer, L. B., Chuwah, C., Ganzeveld, L., Hazeleger, W., Hurk, B. van den, van Noije, T., O'Neill, B. and Strengers, B. J.: A comprehensive view on climate change: coupling of earth system and integrated assessment models, Environ. Res. Lett., 7(2), 24012, doi:10.1088/1748-9326/7/2/024012, 2012.

van Vuuren, D. P., Kriegler, E., O'Neill, B. C., Ebi, K. L., Riahi, K., Carter, T. R., Edmonds, J., Hallegatte, S., Kram, T., Mathur, R. and Winkler, H.: A new scenario framework for Climate Change Research: Scenario matrix architecture, Clim.

Change, 122(3), 373–386, doi:10.1007/s10584-013-0906-1, 2014.~~van Vuuren, D. P. and Carter, T. R.: Climate and socio-economic scenarios for climate change research and assessment: reconciling the new with the old, Clim. Change, 122(3), 415–429, doi:10.1007/s10584-013-0974-2, 2014.~~

~~van Vuuren, D. P., Lowe, J., Stehfest, E., Gohar, L., Hof, A. F., Hope, C., Warren, R., Meinshausen, M. and Plattner, G.-K.: How well do integrated assessment models simulate climate change?, Clim. Change, 104(2), 255–285, doi:10.1007/s10584-009-9764-2, 2011.~~

~~Vuuren, D. P. van, Bayer, L. B., Chuwah, C., Ganzeveld, L., Hazeleger, W., Hurk, B. van den, van Noije, T., O'Neill, B. and Strengers, B. J.: A comprehensive view on climate change: coupling of earth system and integrated assessment models, Environ. Res. Lett., 7(2), 24012, doi:10.1088/1748-9326/7/2/024012, 2012.~~

~~van Vuuren, D. P., Kriegler, E., O'Neill, B. C., Ebi, K. L., Riahi, K., Carter, T. R., Edmonds, J., Hallegatte, S., Kram, T., Mathur, R. and Winkler, H.: A new scenario framework for Climate Change Research: Scenario matrix architecture, Clim. Change, 122(3), 373–386, doi:10.1007/s10584-013-0906-1, 2014.~~

Wang, C., Zhang, L., Lee, S.-K., Wu, L. and Mechoso, C. R.: A global perspective on CMIP5 climate model biases, Nat. Clim. Chang., 4(3), 201–205, doi:10.1038/nclimate2118, 2014.

Washington, W. M., Buja, L. and Craig, A.: The computational future for climate and Earth system models: on the path to petaflop and beyond, Philos. Trans. R. Soc. London A Math. Phys. Eng. Sci., 367(1890) [online] Available from: http://rsta.royalsocietypublishing.org/content/367/1890/833 (Accessed 18 June 2017), 2009.

Williams, P. D.: Modelling climate change: the role of unresolved processes, Philos. Trans. R. Soc. A Math. Phys. Eng. Sci., 363(1837), 2931–2946, doi:10.1098/rsta.2005.1676, 2005.

Williamson, D. L.: Description of NCAR Community Climate Model (CCM0B), Tech. Rep. NCAR/TN-210+STR, 88, 1983.

Xie, P. and Arkin, P.: Global precipitation : A 17-year monthly analysis based on gauge observations, satellite estimates and numerical model outputs, Bull. Am. Meteorol. Soc., 78, 2539–2558, doi:10.1175/1520-0477(1997)078<2539:GPAYMA>2.0.CO;2, 1997.

Yeh, S.-W., Kug, J.-S., Dewitte, B., Kwon, M.-H., Kirtman, B. P. and Jin, F.-F.: El Niño in a changing climate, Nature, 461(7263), 511–514, doi:10.1038/nature08316, 2009.

Yuan, W., Yu, R., Zhang, M., Lin, W., Li, J. and Fu, Y.: Diurnal cycle of summer precipitation over subtropical East Asia in CAM5, J. Clim., 26(10), 3159–3172, doi:10.1175/JCLI-D-12-00119.1, 2013.

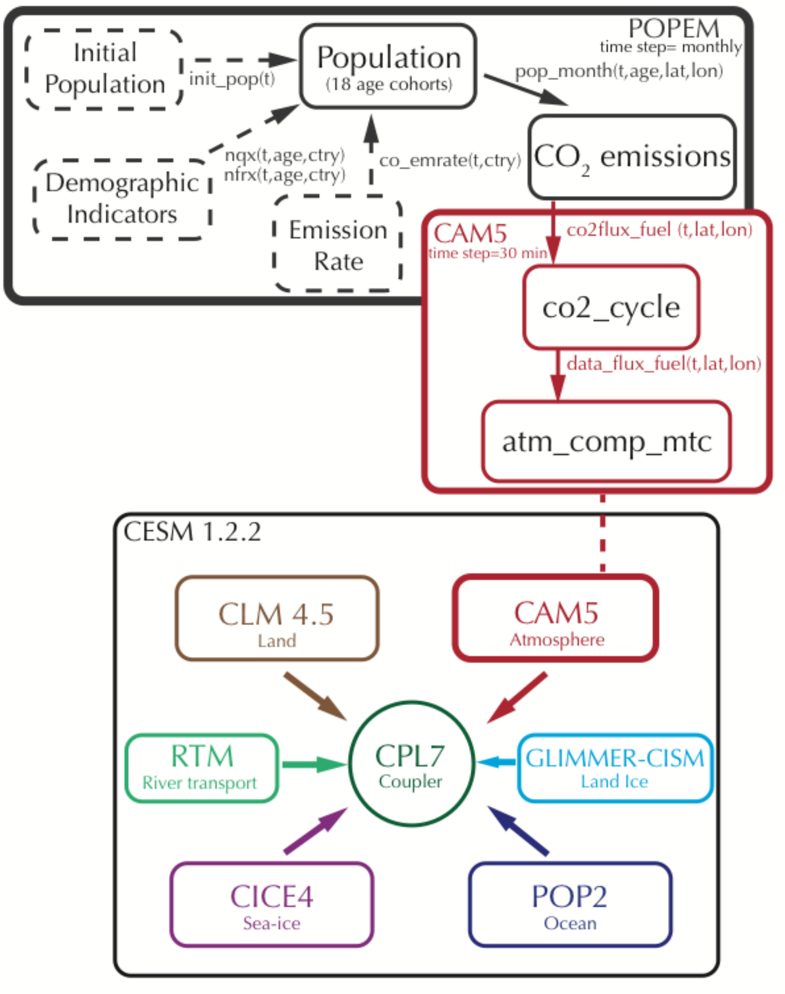


Figure 1: Conceptual schema of the POPEM module coupled with the CAM5 atmosphere module. POPEM requires three input data sets to compute emissions (black dashed rectangles): initial population distribution; demographic parameters (age structure, death, and birth rates); and per capita emission rates by country. POPEM provides a 3D array (time, latitude, longitude) with emissions that are read by the *CO2_cycle* module and passed to the *atm_comp_mct* module which computes the total amount of CO_2 in the atmosphere.

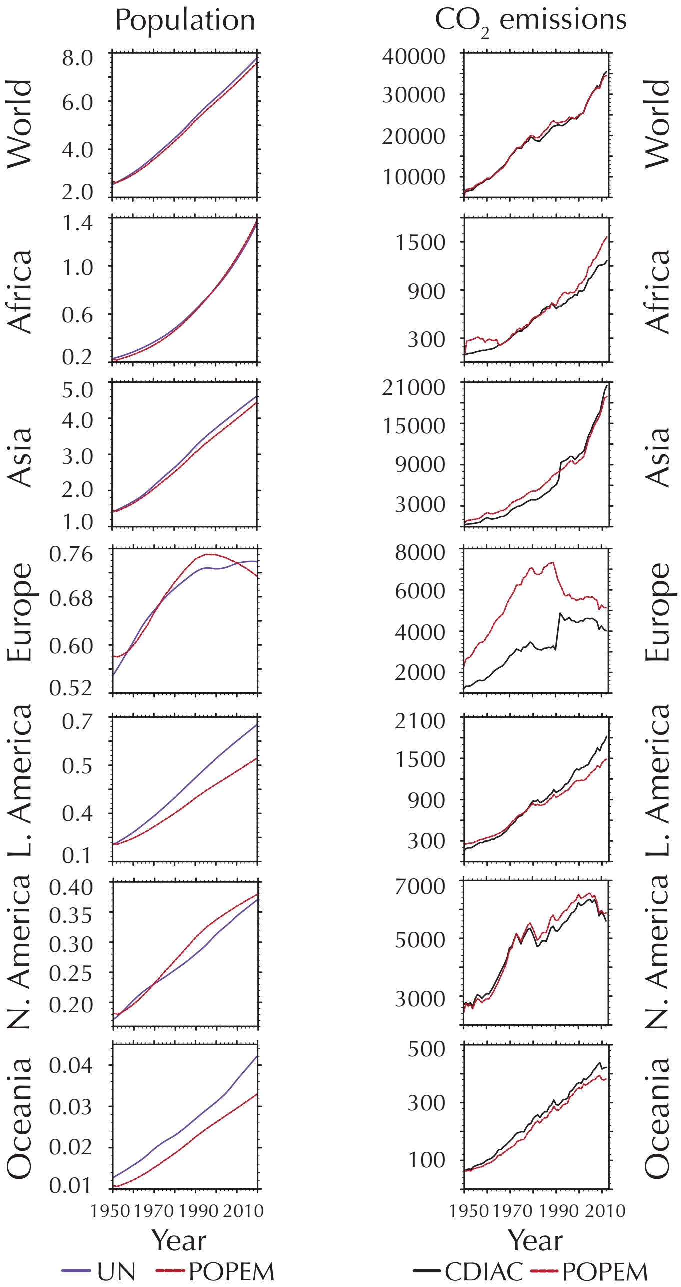


Figure 2: Comparison of the population estimates for the years 1950-2020 (left column) and the historical CO_2 emissions estimates for the years 1950-2012 (right column). The first row compares global data, the second to seventh compare regional data (Africa, Europe, Latin America, North America and Oceania). In the left-hand column, the red line shows the estimates given using POPEM and blue indicates UN estimates. Values are given in thousand millions of people. On the right, the red line shows the estimates given using POPEM and the black indicates CDIAC estimates. Units are given in million metric tons.

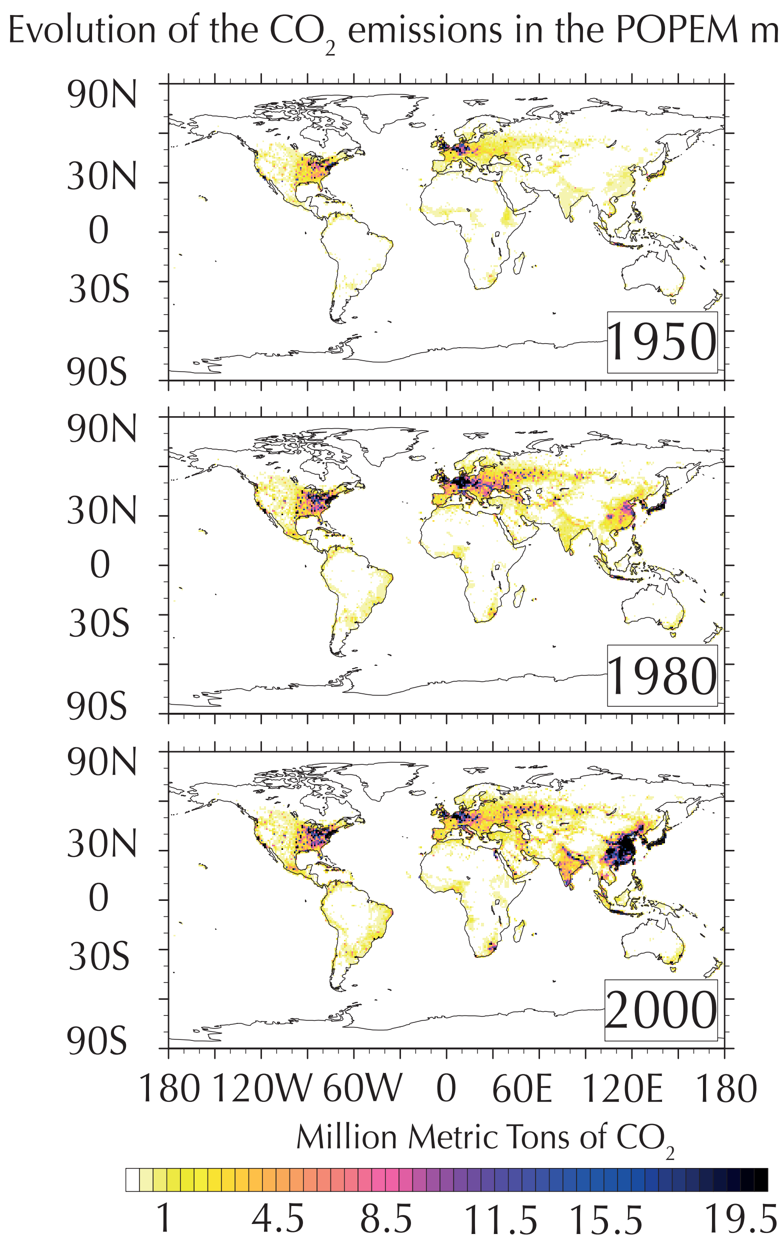


Figure 3: POPEM CO_2 emissions estimates for 1950, 1980 and 2000. POPEM produces a gridded representation of anthropogenic CO_2 emissions using population dynamics and country per capita emissions derived from the CDIAC database. Values are given in millions of metric tons per year.

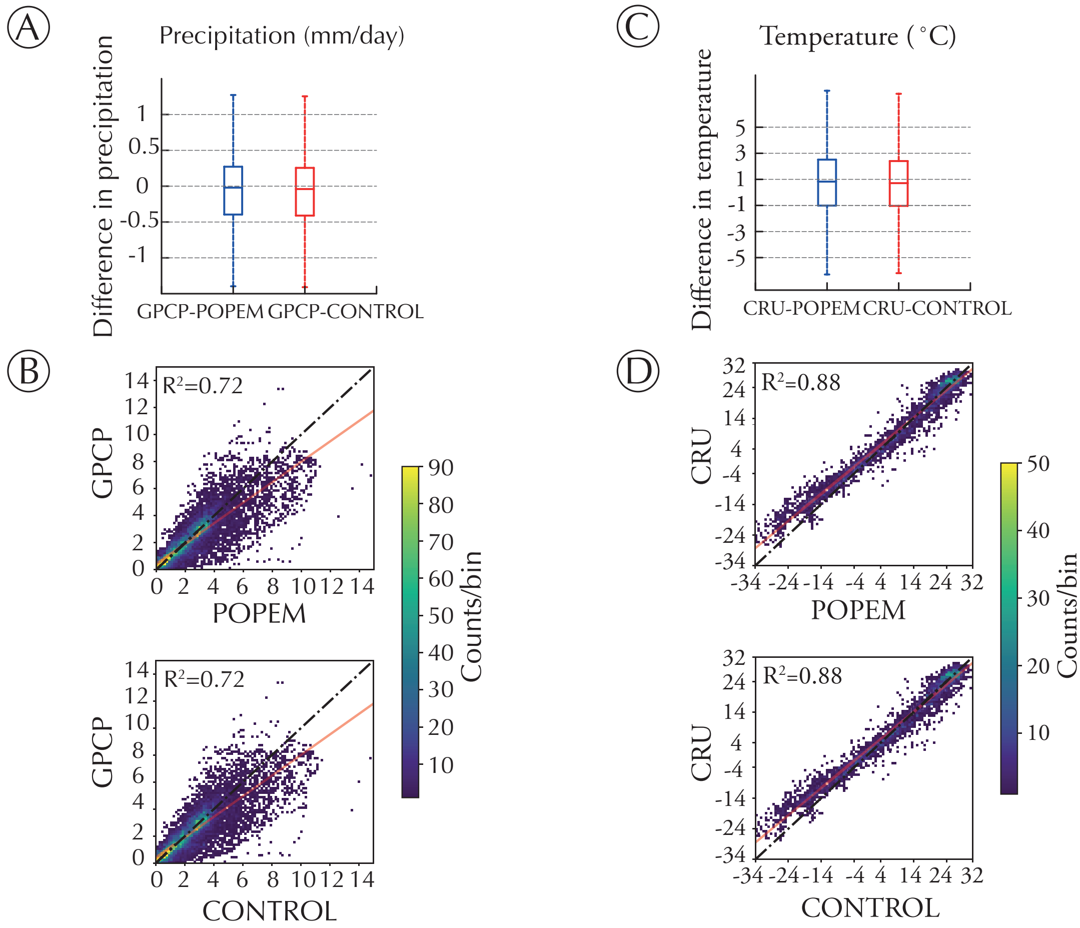


Figure 4: Boxplots of CESM simulation bias for precipitation (A) and temperature (C). (B) Scatter plots comparing the annual mean precipitation (1980-2000) at every grid point for GPCP and CESM simulations (POPEM and CONTROL). (D) Scatter plots comparing the annual mean temperature at every grid point for CRU and CESM simulations (POPEM and CONTROL). Units are in mm/day (precipitation) and in degrees Celsius (temperature).

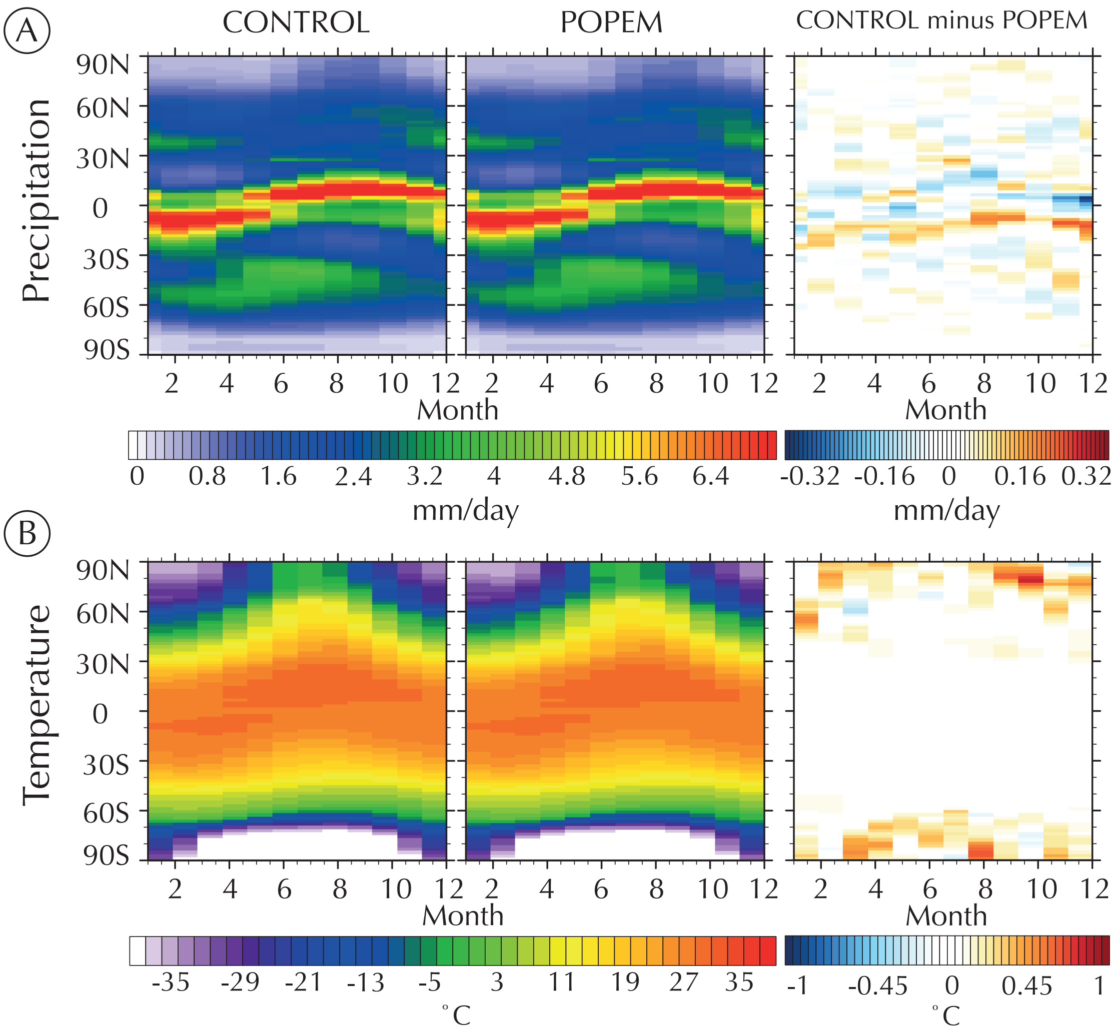


Figure 5: Latitude vs time plots for precipitation (A) and surface temperature (B). For absolute difference graphs, blue represents higher values in POPEM and red represents higher values in the CONTROL. Units are in mm/day for precipitation and in Celsius for temperature.

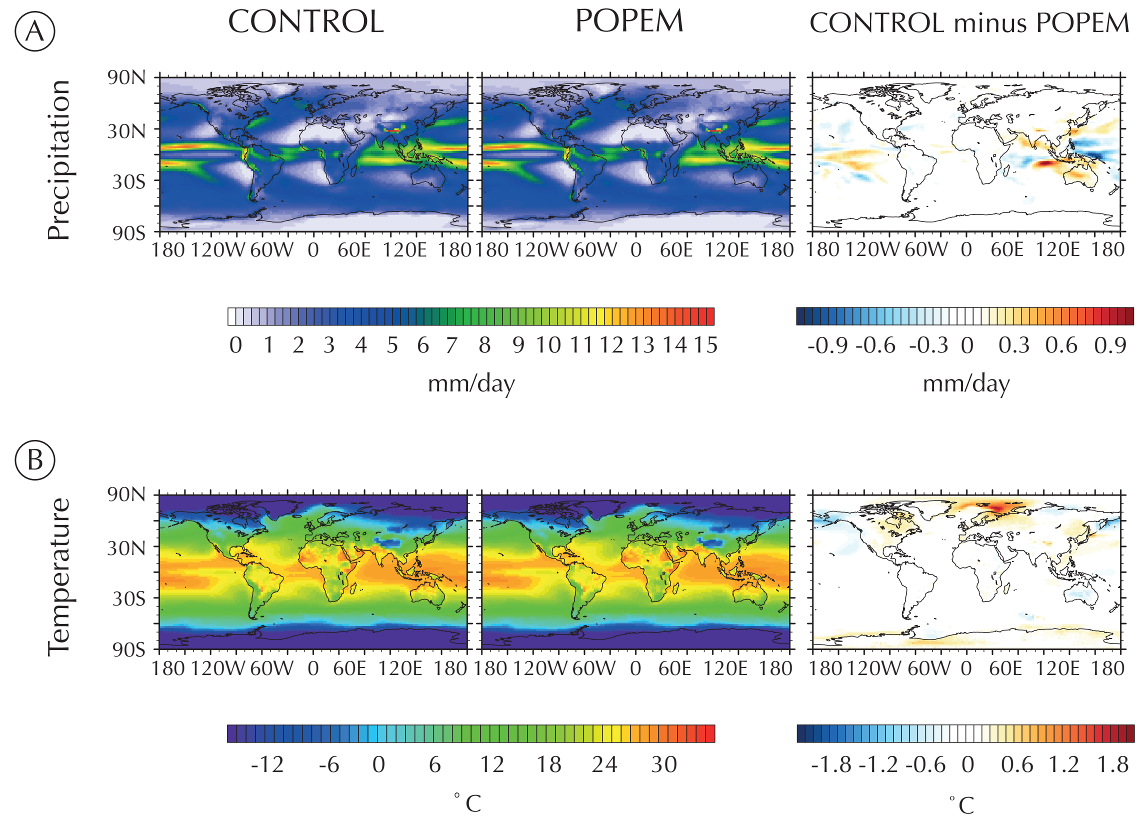


Figure 6: A comparison of global annual mean precipitation (1950-2000) for the CONTROL and POPEM (A). (B) is a comparison of annual mean surface temperatures. The maps in the right-hand column show the absolute differences between the simulations (CONTROL minus POPEM). In these, blues represent higher values in POPEM and reds represent higher values in the CONTROL.

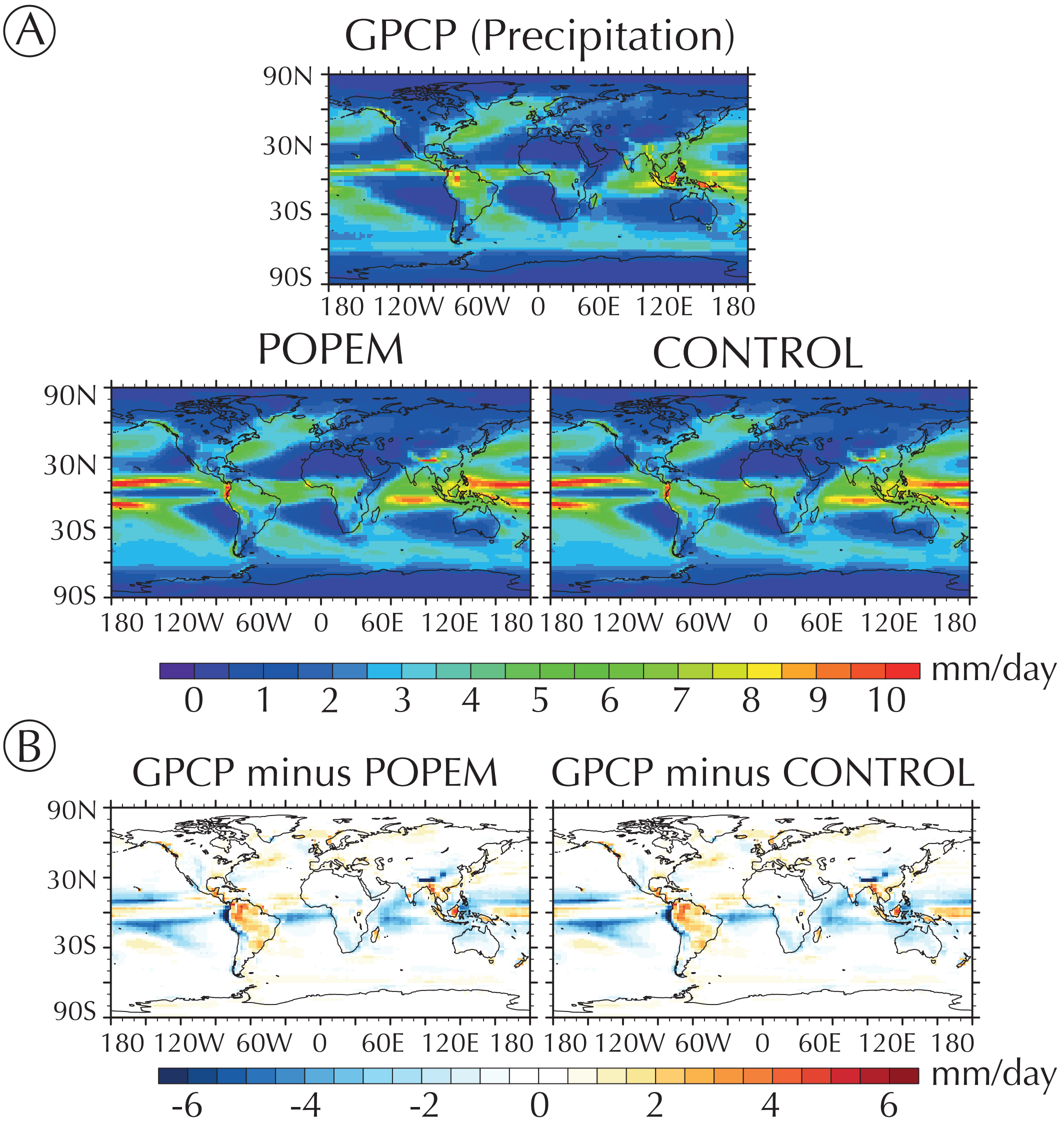


Figure 7: A comparison of the global annual mean precipitation (1980-2000) as simulated by the CESM (POPEM and CONTROL) model and GPCP observational database. (A) Global annual mean precipitation maps for GPCP, POPEM and Control. (B) Absolute difference maps. Units are in mm/day.

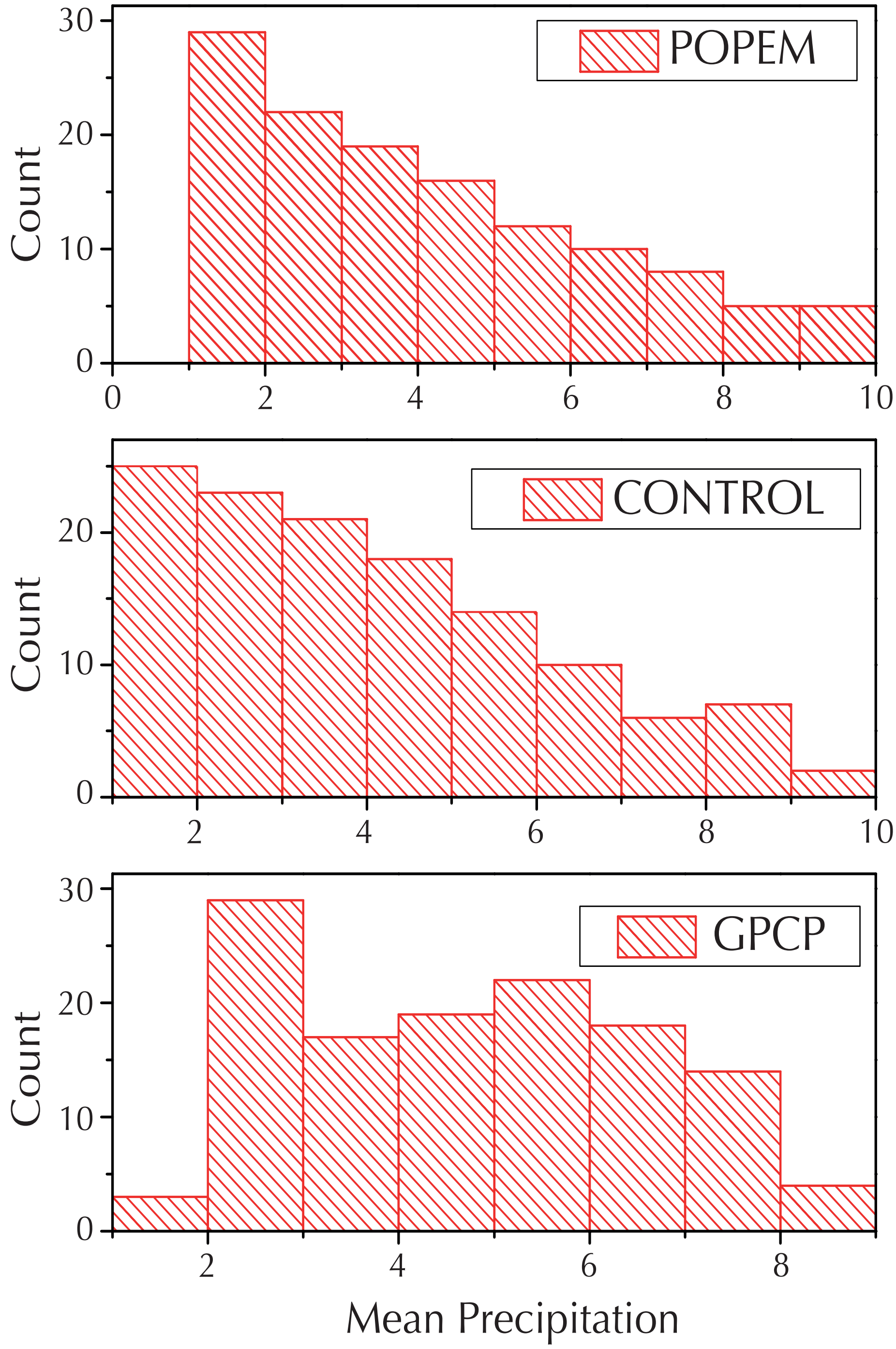


Figure 8: Histograms of the mean precipitation in the El Niño-4 area (5N-5S, 160E-150W) using the POPEM parameterization (top), the standard forcing approach (CONTROL, middle), and the reference GPCP estimates (bottom).

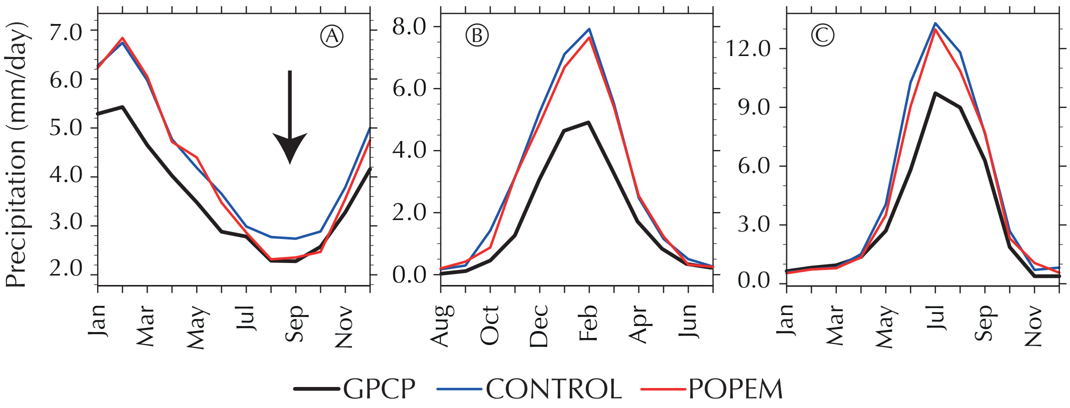


Figure 9: Monthly precipitation (1980-1999) based on GPCP, CONTROL and POPEM for three of the regions with important biases in CESM. (A) shows precipitation for the area affected by the double-ITCZ bias in the Southern Hemisphere (20S-0, 80E-100W); (B) for Australia Top End (30S-10S, 128E-140E); and (C) for the Tibetan Plateau (22N-32N, 78W-92W). The black line represents observations (GPCP), the blue line is the CONTROL case, and the red line is the POPEM case. Units are in mm/day. The arrow indicates the improvement of the POPEM model.

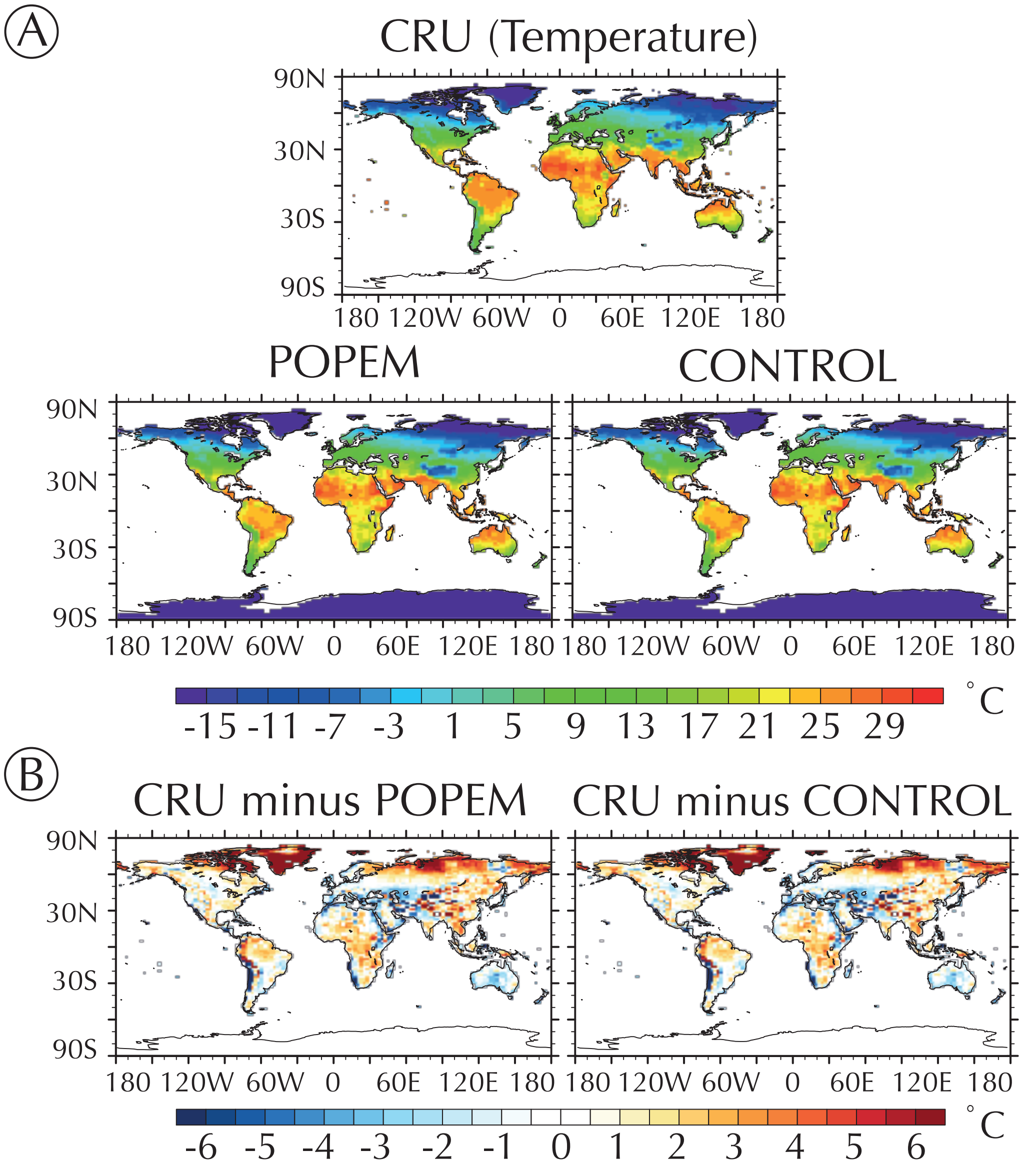
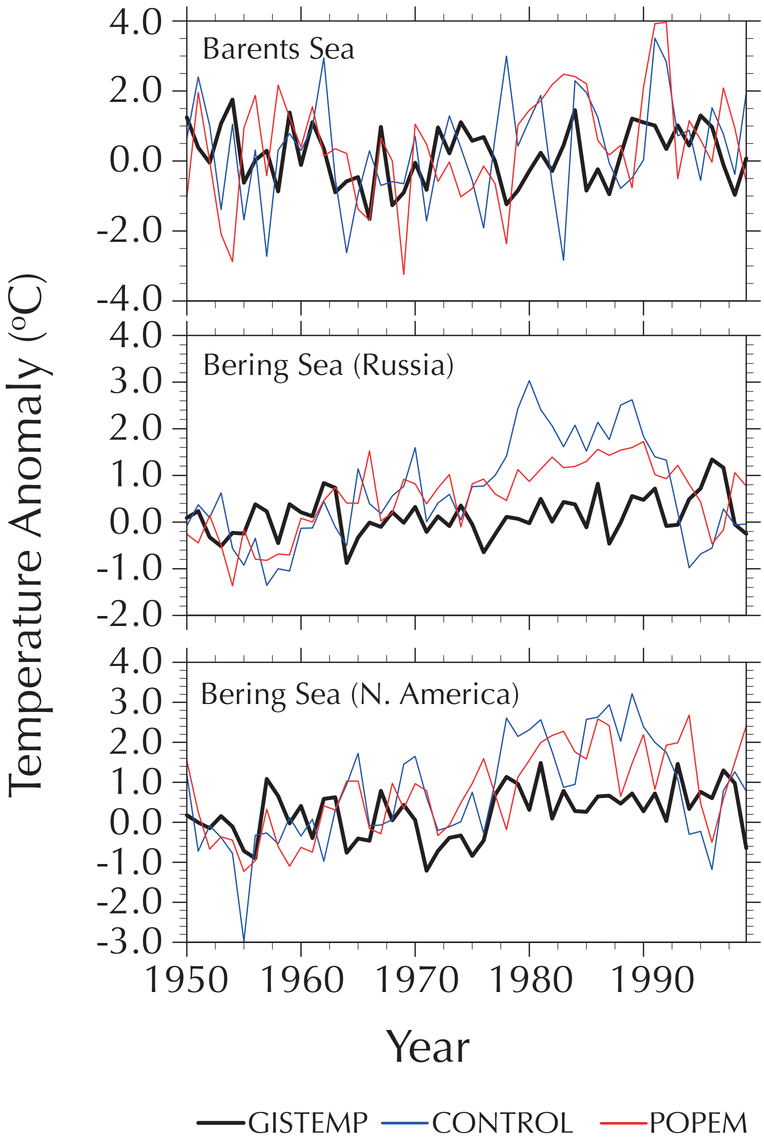


Figure 10: A comparison of the annual mean temperature (1950-2000) as simulated by the CESM model (POPEM and CONTROL) and CRU observational database. (A) Global annual mean temperature maps for CRU, POPEM and CONTROL. (B) Absolute difference maps. Units are in degrees Celsius.



5 **Figure 11:** **A comparison of the annual mean surface temperature anomaly between GISTEMP, CONTROL and POPEM from 1950 to 1999. (Top) represents the Barents Sea (68N-80N, 19E-68E); (middle) Russian part of the Bering Sea (50N-65N, 150E-180E); and (bottom) American part of the Bering Sea (50N-75N, 140W-180W). The black line represents observational data (GISTEMP), the blue line is the CONTROL case, and the red is the POPEM case. Anomaly was referenced to 1951-1980 period.**

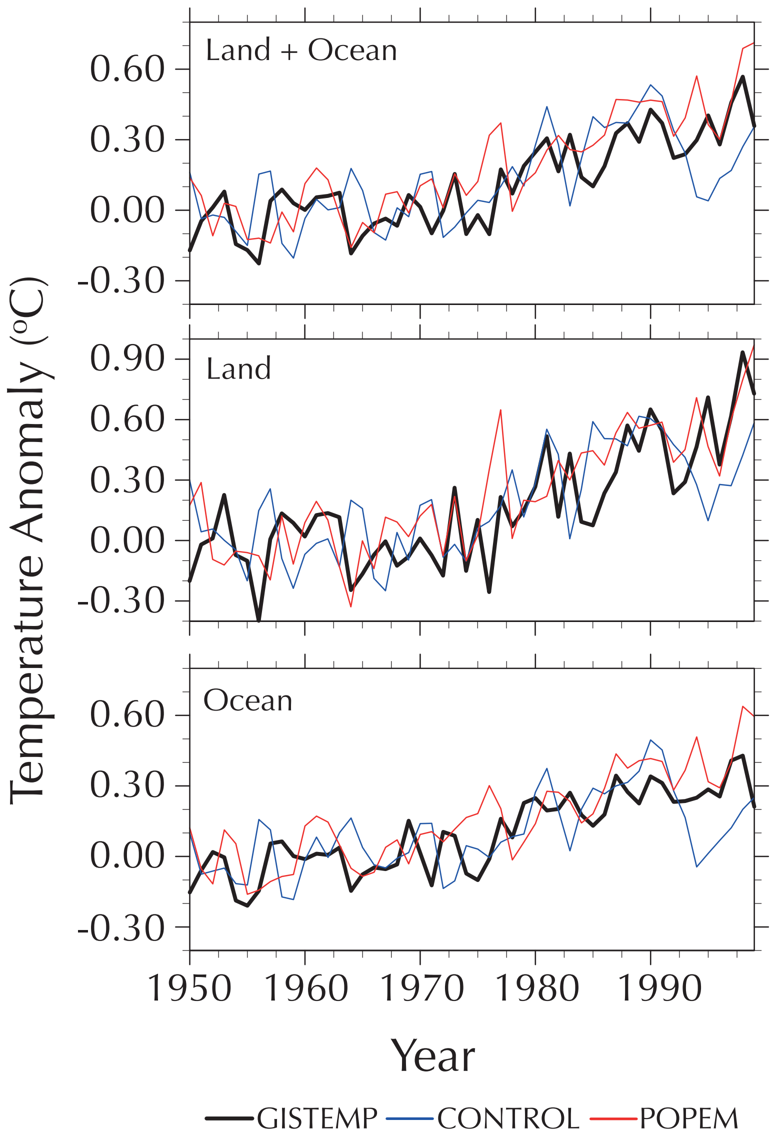


Figure 12: A comparison of the global annual mean surface temperature anomaly between GISTEMP, CONTROL, and POPEM from 1950 to 1999. (Top) global; (middle) land; and (bottom) ocean. The black line represents observational data (GISTEMP), the blue line is the CONTROL case, and the red is the POPEM case. Anomaly was referenced to 1951-1980 period.

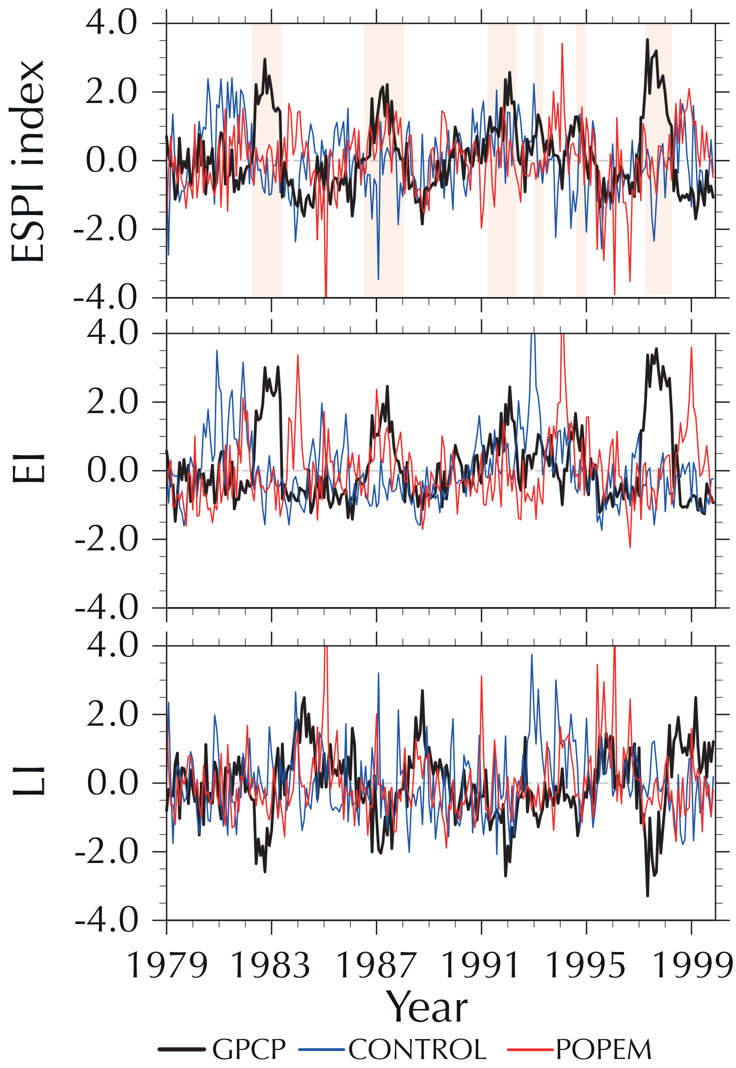


Figure 13: Time-series of precipitation anomalies for the ENSO region after Curtis and Adler (2000). (Top) ENSO Precipitation Index (ESPI); (Middle) El Niño Index (EI); and (Bottom) La Niña Index (LI). The Black line shows GPCP data, the blue line is the CONTROL case, and the red line is the POPEM case. Orange shading denotes El Niño years defined as consecutive months (minimum 3) with NIÑO3.4 sea surface temperature anomalies (5N–5S, 170–120W) greater than +0.5°C.

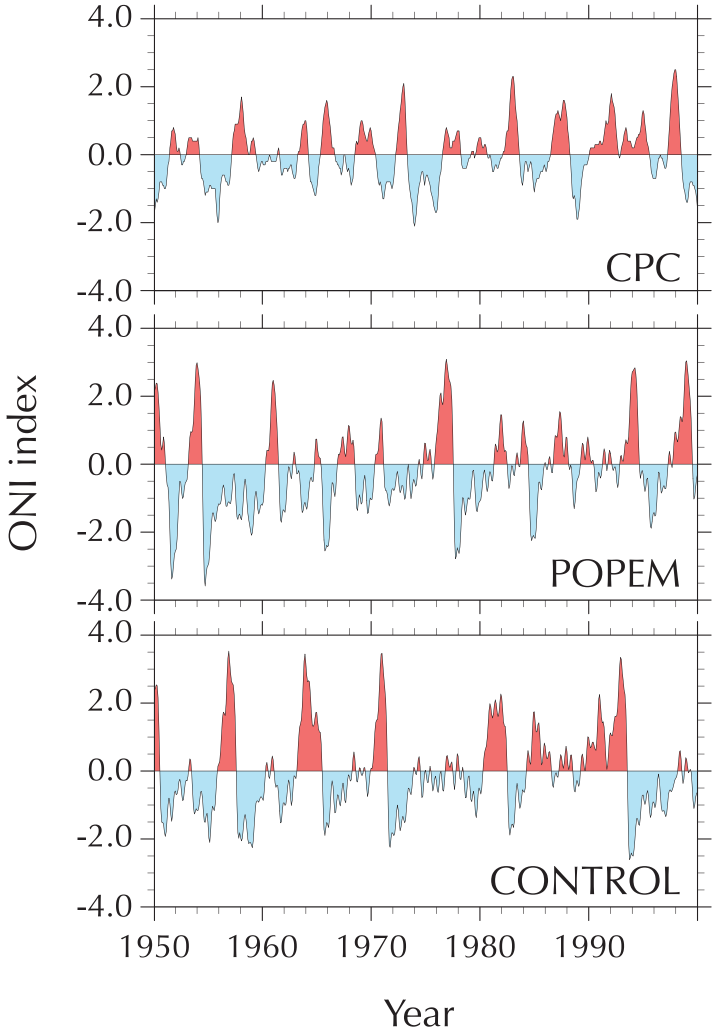


Figure 14: Comparison of the Oceanic el Niño Index (ONI) for CPC (top), POPEM (middle), and CONTROL (bottom) cases. El Niño and La Niña are defined according to Kousky and Higgins (2007): 3-month running mean with anomalies greater than +0.5°C (or-0.5°C) for at least five consecutive months in NIÑO3.4 region. The base period for computing SST departures is 1971–1999.

Table 1. Comparison of the ONI index for the period 1950-1999. The table compares the ability of the models to reproduce the number, strength, and duration of el Niño events.

Source	Number of events	Agreement [1]	Disagreement [2]	Intensity Bias $_{avg}$ [3]	Duration $_{avg}$ [4]
CPC	14				10.3
CONTROL	7	33	121	0.59 ^{o} C	19.4
POPEM	10	37	121	0.22 ^{o} C	11.4

[1] The number of months that CPC and CESM agree on El Niño.[2] Disagreement defined as the number of months where CPC and CESM obtain opposite results. [3] Intensity: (|CESM ONI| – |CPC ONI|)/number of cases (units in degrees Celsius). [4] Mean duration of El Niño event (in months).