The purpose of this paper is to document the GCMs derived from HadCM3 that have been and will continue to be used by the authors for a wide range of climate research activities. This is an appropriate purpose for a GMD paper. The paper is clearly and well-written and serves the purpose. Its comparison of the results from various HadCM3-related configurations with each other and with CMIP5 models, to give an overview, is particularly useful. Although HadCM3 is quite an old model, it is still important. In the Web of Science, the paper of Gordon et al. (2000) on HadCM3 has been cited at least 100 times per year in the last decade. I have first some general comments, then some specific ones. I'm sure that my comments can be addressed, if the editor agrees that they should be, but since they're quite extensive I'm classifying them as implying major revisions.

We thank the review for the time they have taken to improve the paper with their detailed comments.

General comments

(1a) It's rather unusual for a paper to document a model that was mostly not developed by the authors, especially when other papers already exist about the model. Moreover the authors are not the sole users of this model. Of course the authors are aware of these points, and have cited papers other than theirs; I have suggested some more citations, and some minor rephrasing to avoid misinterpretation. In a way, this situation is like an open-source development (these models aren't open-source, although readily available).

We do agree that the paper is unusual in that we have developed relatively small amount of the overall code, but it arose because reviewers were increasingly critical of the lack of documentation of the specific model versions that we were using in many other papers. In addition, reviewers were also critical of the overall skill of our models simply because it was old. Hence we wished to show that, at least for its mean climate, it remains competitive with the state-of-the-art models.

To better reflect some of these issues, we have invited Pope and Gordon to be co-authors and we am pleased to say they have accepted. Moreover, Michel Crucifix assembled the updates for MOSES2.1 and is also now an co-author.

(1b) In view of this, the paper should focus on the aspects which BRIDGE specifically has contributed. For example, while the summaries of sects 2.1-2.4 of HadCM3 are clearly written, I feel that this much information is not needed, because it's essentially documentation of the original HadCM3, for which Pope et al., Gordon et al., and references therein should be cited. For the purpose of this paper, sect 2.5 is the important part, giving the differences of the BRIDGE HadCM3-M1 from the model of Gordon et al. These are stated to have very little effect.

We have changed the tone to ensure there is no confusion about who wrote the model. The first sections are a brief summary of the original model for completeness. We feel that it does help the reader understand the context of our paper, though we now strongly emphasis that the “new” aspects start at section 2.5.

(2) The nomenclature of the models could be confusing. The purpose of this paper is to describe models used at BRIDGE, but these have to be clearly related to or distinguished from other models, in the same family, that have been used at the Met Office and elsewhere. HadCM3 and HadAM3 are models developed by the Hadley Centre.

On p3 line 12-14, the authors write, "Since its introduction, HadCM3 (and related models) has undergone a substantial number of changes, bug fixes and adaptations, such that few of the versions of the model used now are truly identical to their original model description in Pope et al. (2000) and Gordon et al. (2000)." I think this statement is debatable. It is likely that Hadley Centre users of HadCM3, and at least some other users, would say that it is not *really* HadCM3 if it’s different from the model described by the papers of 2000, as far as the AOGCM is concerned, except for bug-fixes and mods needed for porting. That is, HadCM3 has not been revised. Of course it’s hard to decide what’s a bug-fix versus a scientific change - that’s why this is debatable.

It turns out that the BRIDGE HadCM3-M1 is practically identical to the model of Gordon et al. (2000) (section 2.5). This is useful to know, but I’d suggest that it means it would be clearer to call this model just "HadCM3" throughout the paper.

We agree that some of the models used by BRIDGE are very similar to those presented by the Met Office. We have further clarified this at a number of points in the text. However we believe that the changes made by BRIDGE should be defined and described. This removes ambiguity surrounding HadCM3 and the variants used at Bristol which is the key aim of the GMD journal (and has been asked for by reviewers of our papers).

Moreover, there is already some ambiguity in the use of the name HadCM3 throughout the community, particularly related to the land surface scheme. In one sense, the reviewer is correct to say that the core AOCGM “defines” HadCM3 but it is now common to use alternative land surface schemes and, in most cases, these are still referred to as HadCM3. For instance, it is difficult to determine if the HadCM3 featured in the model evaluation figure 9.1 in the IPCC AR5 report (corresponding to our figure 1) was the original model, or included MOSES2.1/TRIFFID or MOSES2.2/TRIFFID. Most of the text implied it was the original model but the model description table (9.A.1) references the Cox et al TRIFFID description paper (which was using MOSES2.1) and a Mercado et al paper which is MOSES2.2!

So we argue that it is important to improve the nomenclature for the model variants described here to more clearly distinguish the different versions as well as those developed at the Met Office. We have changed the nomenclature of the model acronyms to include the letter **B** for Bristol, so HadCM3-M2.1 becomes HadCM3B-M2.1 etc. We have continued to use the HadCM3 acronym despite the differences in the model to that described by Gordon et al. and Pope et al., as the vast majority of the code remains the same.

To more clearly highlight what changes that have been made to the Bristol variants, we have moved the bug fixes section to section 2.1 before the descriptions of the components and split this section to define changes made to HadCM3-M1 and HadAM3.

In all cases we have retained the MOSES suffixes, as the name HadCM3 appears to frequently be used as a generic term regardless of the land surface scheme used. We think it would be a useful precedent to be more specific about which version of HadCM3 is being used to clearly distinguish this version from others using MOSES2.1 or 2.2.

However, we learn that the BRIDGE HadCM3L is a very different model from the earlier Met Office model HadCM3L (p14 line 32). (A reference should be given for the Met Office HadCM3L - I think it might have been first used though not named by Cox et al., Nature, 2000). Similarly, the BRIDGE HadAM3H is said to be different from HadAM3H as used by the Met Office (p14 line 24-25). It is therefore confusing to call the Met Office and BRIDGE models by the same names.

Agreed. We have now resolved this naming confusion by adopting the HadCM3B nomenclature throughout where the model departs from the Met Office model.

The Cox reference has now been added at the first mention of HadCM3L who as you say first uses the L flavour of the model (but does not call it HadCM3L).

At least in the context of this paper, I suggest that models from BRIDGE and the Met Office are given different names, so that the reader can be clear, when differences from HadCM3 are mentioned, which are the ones introduced at BRIDGE, and which are the same as in the Met Office version. For HadRM3, is Hudson and Jones (a technical report on a particular application) the only reference that can be cited?

The reference to Hudson and Jones was for HadAM3H. We have added a citation for HadRM3 too to help. We have also added a further reference for the use of HadAM3H.

There are several published versions of FAMOUS. It would be useful if the authors could relate the BRIDGE version to those documented by Jones et al. (2005), Smith et al. (2008) and Smith (10.5194/gmd-5-269-2012, 2012). The Smith papers identify versions of FAMOUS by the run ID of the definitive UM basis file, so it would be useful to relate the runs detailed in Sect 7 to theirs. Also, FAMOUS is documented at www.famous.ac.uk, which could be cited.

Our version of FAMOUS is identical to that described in the papers above. We have clarified this in our text. References to FAMOUS are only included to complete the set of models based on HadCM3. We have made that clearer in our text.

(3) Effect of differences in configuration. It seems to me that there's not a very close connection between the model documentation of sect 2-4 and the results of sect 5. It would be valuable, wherever possible, to relate the differences in results to the differences in formulation. I appreciate that some of this is done, and that it's hard, but whatever more can be done would be useful, since it's the kind of information that would be relevant to deciding or understanding the choice of model for a particular purpose.

We've added discussion in the temperature and land sections of section 5 about how the differences in sections 2-4 affect the performance of the model.

(4) Computational cost (i.e. speed) is an important consideration in the choice of model. If one had infinite computing resource, for example, one would probably not prefer FAMOUS to HadCM3. It would be informative to collect some numbers for the relative computational cost of the model e.g. in Table 1.

This is a good point raised by both reviewers. We have included a new table of run speeds for each of the MOSES1 models in typical configurations, run on the same machine (table 3).

Specific comments

p1 line 8. I would delete "originally". Subsequent developments are smaller changes than the original development; HadCM3 is still essentially the same.

We agree, this was misleading and has now been removed.

p1 line 9-10. "but is now largely being replaced by more recent models" strikes me as an odd thing to say. It's true that HadCM3 is now little used by the Hadley Centre, which mostly uses later Met Office models. Some other centres, including BRIDGE, use HadCM3, but these centres, including BRIDGE, use other and later models too. Most centres, including the Hadley Centre, routinely use a range of models, older and newer, for different purposes. For reasons the paper explains, there are purposes for which older models are more suitable and cannot be replaced by later ones. I would suggest deleting this phrase.

We would like to retain this phrase because HadCM3 has for many of the areas for which it was originally developed, been effectively completely superseded by more recently released models. These newer models either have higher resolution, include more complete physical/biological/chemical representations or have differences in the fundamental aspects, such as dynamics or clouds etc. We have thus rephrased this:
"... has been heavily used during the last 15 years for a range of future (and past) climate change studies but has now been largely superseded for many scientific studies by more recently developed models."

p1 line 10. This paper is about the models used at BRIDGE, but other places use HadCM3 in the Met Office version. This sentence and the next two could seem to imply that BRIDGE is responsible for development of these models. To avoid that implication, you could slightly rephrase e.g. It continues to be used by various institutions, including by the BRIDGE (Bristol Research Initiative for the Dynamic Global Environment) research group at the University of Bristol, who have made adaptations over time to the base HadCM3 model. These adaptations mean that the original documentation is not entirely representative of the models used at BRIDGE, where several other configurations are in use which now differ from the originally described model versions.

This is a very reasonable point and the sentence in question has been adjusted.

p1 line 17. In the title and here you mention "version 1.0" for this suite of models, but you don't mention it again in the paper, which describes each of the configurations individually. Is it useful to assign a version number to the entire suite? If it is, it would be worth explaining what defines a version of the suite, and what would qualify as a new version.

The GMD guidelines require that all model description papers have a version number. This leaves open the possibility of publishing updates in the future. When this version number is changed, and to what (e.g. v2.0, v1.1), will depend on the modification(s) being made and will be the topic of future discussions.

p2 line 4. It would be worth spelling out that "model" is a general circulation model. Not all climate models are GCMs.

We've added: "(all of which can be classed as General Circulation Models, GCMs)".

p2 line 6. The statement is correct that the HadCM3 family is in use at BRIDGE, but it could seem to imply that BRIDGE is the only group which uses HadCM3 etc. still. As the authors know, and have stated at p3 line 14, it has other active users too.

Entirely fair point, and we have modified the text to reflect that we are not the only users.

p2 line 7. While it's true that Gordon et al. (2000) describe developments to the ocean model, I would say that main achievement and interest of that paper is HadCM3 i.e. the coupled model. Thus I would suggest phrasing this slightly differently e.g. "HadAM3 (Pope et al., 2000). Together with improvements to the ocean GCM, this enabled the development of HadCM3 (Gordon et al., 2000), which was one of the first coupled atmosphere-ocean GCMs that did not require ...".

The text has been edited as per the reviewer suggestion.

p2 line 9. Replace "was" with "has been", since it's still in use. It has been used for the past as well as the future. I think that at this point it would be appropriate to cite

some of the major uses of HadCM3 that don’t come up later, such as detection and attribution (Stott et al., Science, 2000), unforced variability (Collins et al., Clim Dyn, 2001), climate projection (Johns et al., Clim Dyn, 2003), uncertainty and constraints on projections (Stott and Kettleborough, Nature, 2002), decadal prediction (Smith et al., Science, 2007). I’ve made other suggestions later too.

“Was” replaced. We have also inserted the suggested references.

p2 line 10. Why is "though" at the end of the sentence?

Apologies, it seems like the end of the sentence was cut off and it was somehow missed in the editing process. This has been corrected.

p2 line 16. Reichler and Kim (BAMS, 2008) could be cited here as well.

Added.

p2 line 17. UKESM hasn’t been introduced.

Both reviewers pointed out the unexplained acronym. The text has been updated to say UK Met Office Unified Model (UM) rather than specifically UKESM.

p2 line 24. Hewitt et al. (GRL, 2001) could be cited here as the first use of HadCM3 for a snapshot of palaeoclimate.

Added.

p2 line 26. Faster models are also useful for investigation of anthropogenic change on long timescales; HadCM3 has been used for that as well e.g. Gregory et al. (GRL, 2004), Ridley et al. (Clim Dyn, 2005).

We have added this to the text and cited these examples.

p2 line 25. Although multi-centennial rather than multi-millennial, I think it would be relevant to cite Tett et al. (Clim Dyn, 2007) as the first use of HadCM3 for a study of long-term past climate change. Among examples of more recent uses of HadCM3 for such a purpose, I would say that Schurer et al. (Nature Geosci, 2013) could be cited.

We have added Tett et al., as an early multi-centennial example here.

p2 line 27-29. What boundary conditions are meant here? I usually understand BCs as concerning forced climate change. Of course there are numerous earlier examples of climate-change experiments with HadCM3. Large initial-condition ensembles are a good application too, I agree, but I would say that a greater use of HadCM3 has been for parameter perturbation ensembles, the next point in this para, so maybe the order of these should be reversed.

We have provided clarification of the meaning of the term boundary conditions and have reversed the order of typical usage of the models.

p2 line 29-31. I'm unclear what distinction is being drawn between "multiple runs to explore the sensitivity" and "calculating probability density functions". Could these be merged? Murphy et al. (Nature, 2004, the first QUMP paper), and Stainforth et al. (Nature, 2005, the climateprediction.net paper) should be cited.

These two have been merged together and Stainforth et al. (2005) cited.

p3 line 9. Jones et al. (Clim Dyn, 2005) should be cited here as well, since it was the first publication of the FAMOUS AOGCM, and Smith (2012, 10.5194/gmd-5-269-2012), which documents the most recent versions.

We have added the additional citations as requested.

p3 line 12-14. See general comment (2), on nomenclature. Since this paper doesn't intend to survey all users and variants of HadCM3, it might be better to move the statements in the latter part of this para, from line 17, to the start of the para, and thus begin by stating that BRIDGE has modified HadCM3 in various ways.

We have kept the structure but modified the emphasis that we are only describing BRIDGE usage. Moreover, we also correcting for the lack of documentation of MOSES2.1, and HAdCM3L which has been used by many groups not just BRIDGE.

p3 line 29. At line 28 on p4, the authors explain that HadCM3-M1 is not exactly the model of Gordon et al. (2000). It would be useful to say that here, since I assumed that's what it was on first reading. See also general comment (2), on nomenclature.

This has now been made clearer earlier in the text with the specific section to these modifications highlighted/linked.

p5 Table 1. Please give units for all quantities that aren't dimensionless. Some of these parameters are generally understandable, such as resolution and "no Iceland", but I suspect that many or most of them are not self-explanatory. While I appreciate the value in documentation, I think we have to consider what the purpose of the paper is. As far as most readers are concerned, the interest in these parameters will be their physical effect, and the consequences of changing them; if they are of sufficient interest to mention, these should be described in the text, with references to the literature of the schemes where appropriate. Writing down the values which appear in the configuration of these schemes is technical documentation, rather than scientific, so perhaps it's not needed in the paper, or maybe it would belong better in an appendix.

We have added the missing units to this table. We have elected to retain this table as it gives a broad easy reference comparison of the models in terms of resolution and inclusion or not of specific schemes (e.g. vertical ocean diffusion). Also the model parameter values given in this table are referred to in several places in the text, and are therefore integral to the detailed description of the different configurations that we are aiming for.

p6 line 8. This is not an accurate statement. The base code of UM4.5 is available at that URL, but the scientific definition of HadCM3 also depends on a lot of code mods. This is explained in Sect 7, but should be mentioned here too, or you could delete this statement and refer to Sect 7.

It has been clarified that extra modifications are required to the base code, along with a reference to sect. 7 where this is explained.

p14 line 32. See general comment (2) on nomenclature re HadCM3L.

This sentence has been moved to the first paragraph in Section 4.1 (second sentence), and the model name has been updated according to the reply to comment (2).

p15 line 10. I'm not sure it's as simple as that. I think the introduction of the GM scheme is important in alleviating the need for flux adjustment.

We have removed the reference to HadCM2, and instead reported the findings of Jones et al: "Jones et al (2003) investigated potential modifications to allow increased heat transport through this region, thus alleviating the unrealistic buildup of sea ice in the Nordic Sea, and concluded that the removal of Iceland was the preferred solution. With this modification, the improved meridional overturning circulation leads to more realistic heat transports in the coupled system and alleviates the need for flux correction."

p15 line 16 and Table 1. The coefficient being described is the vertical diffusivity - I would not call that a mixing rate. HadCM3 has a prescribed depth-dependent background vertical diffusivity and a Richardson-number-dependent part, which is important near the surface. In FAMOUS, the latter part is omitted. The mixed-layer scheme is distinct from the vertical diffusion scheme.

Table 1 has been updated to describe this coefficient as the "vertical tracer diffusivity" and the option used in BRIDGE versions of HadCM3L and FAMOUS is now described as "constant background value". The text of section 4.1.2. has been updated similarly.

p15 line 25-26. As far as I know, FAMOUS uses the same equation of state (UNESCO) and diffusion scheme as HadCM3 (except as noted in the last point).

We have clarified this in the text. FAMOUS and HadCM3 do use different approximation of the equations of states, but only within the diffusion scheme.

p16 line 24-25. See general comment (2) on nomenclature re HadAM3H.

Please see our previous responses on nomenclature.

p16 line 4 and Table 1. I don’t think the mention of RSOL will be intelligible, unless you explain what this means physically.

This has been updated in the text to briefly explain this term.

p19 line 26. Since it was earlier concluded that the two HadCM3s are the same, maybe it would be clearer to say "the CMIP5 historical experiment of HadCM3 done at the Hadley Centre", to avoid implying it’s a significantly different version. Can you give a reference for this experiment?

The text has been amended as suggested and references added.

p21 Fig 3a caption, same comment, "our version of HadCM3-M1" implies that it differs significantly from the Hadley Centre version.

Text modified.

p29 para from line 21. The net meridional freshwater transport may not be a reliable indicator of bistability. Sijp (Clim Dyn, 10.1007/s00382-011-1249-0, 2012) shows that it’s the derivative wrt AMOC that may be the determining factor. Hawkins et al. (GRL, 2011, 10.1029/2011GL047208) demonstrate AMOC bistability in FAMOUS.

We have stated that freshwater transport may not be a reliable indicator and there remains uncertainty over this hypothesis, citing Sijp, (2012). We have also stated that FAMOUS has been shown to exhibit bistability and included the Hawkins et al. citation.

p30 line 3. Please give a citation to someone it is known by.

This understanding comes from personal communications. To avoid ambiguity, we have removed this to merely state what the models show here.

general comments =============

This paper describes the (largely shared) configurations of a family of climate models as used in a significant research group at Bristol University. The central family member, HadCM3, is well-known and was originally described in papers (and a number of non-peer reviewed technical reports, still readily available) published almost 20 years ago, although the other variants are much less well-described in the available literature, as far as I am aware. Despite the rather elderly nature of the models in question, they are still used and useful, and have been developed and modified in ways that I think sufficiently justifies revisiting and clarifying their documentation in the way the authors have done here. Treating HadCM3 and its spectrum of (roughly) resolution based variants as a family whose commonality and differences are best described together in one paper is a good approach, I feel.

Whilst it is impossible to cover every aspect of the performance of a global climate model, this paper covers a usefully illustrative spread of material, and is uniformly clear and well written. I have a few comments on specific areas, as detailed below, but on the whole I think the paper could be published in largely its current form.

We thank the reviewer for their kind comments and the time they have dedicated to helping us improve this paper.

My main general concern is based in the fact that this is a model description paper with little authorial connection to, or real acknowledgement of, the people who originally developed, coded and made the model available - presumably mostly Met Office and UGAMP/CGAM staff in the 1990s. As far as acknowledgement goes, I'm not sure how best this could be done, but it does feel like a very significant nod should be made in that direction.

As discussed in our response to reviewer 1, we accept that this is an unusual paper since we are effectively documenting small modifications to a large existing model that we did not develop. We have written the paper because a number of reviewers of our other papers have remarked about the lack of details for our modifications of the model. They have also opined that the model is not good enough (compared to CMIP5 models). We therefore decided that this paper is required.

We share your concerns about suitable credit to the original programmers, and therefore before we started writing the paper we contacted the Met Office, and they confirmed that they were happy with us writing the paper. However, we have addressed the reviewers concerns by inviting Vicky Pope and Chris Gordon to be co-authors. We have also invited Michel Crucifix (who brought together most of the HadCM3/MOSES2.1 version that is the centre of our work.

In response to your comments, we have revised throughout the paper to strengthen the fact that we are making modest modifications of a large existing code, and we have added to the acknowledgements to thank everyone involved in the development of the model for their work.

Content-wise, it would be good to be a little clearer on precisely what the BRIDGE authors have themselves added over the MetOffice/CGAM provided code and mods, especially considering that the full documentation for the Gordon et al HadCM3 is still available and that the HadCM3-M1 configuration described here is described as being almost identical to it.

We've gone through the text and added stronger phrasing to clarify what are original model stuff and what are unique to BRIDGE.

We have also added a further comment with respect to HadCM3B-M1. The original UM distribution of the code had a line of code which had a serious compiler specific problem (identified by Dr. Lois Steenman-Clark, whom we acknowledge). On some compilers, it produced (statistically) the same results as shown by the Hadley centre. However, using alternative compilers the model was more than 0.75C cooler in the global mean. Some non-Hadley centre publications of HadCM3-M1 contained this bug. It is thus very important for users to be aware of this. We have added text to clarify this.

specific comments =============

WM="worth mentioning"

page 1, line 17: "version 1.0" - is the intention to upgrade the version number for the whole set of configurations any time there is a bug-fix or change that may only affect one of them? Putting the reference simulation ids in a more prominent place might help with the version/configuration tracking

The GMD guidelines require that all model description papers have a version number. This leaves open the possibility of publishing updates in the future. We are currently making some further modifications and whether these represent new version numbers will be decided when we publish.

p2,l10: "though" seems extraneous

We have removed the word "though" from the end of the sentence.

P2, l17: "UKESM" the acronym should be explained, and a source given for this information. Perhaps "HadGEM3" is meant?

Both reviewers pointed out the unexplained acronym. The text has been updated to say UK Met Office Unified Model (UM) rather than specifically UKESM.

p2,l29: "Roberts et al. in review" - reference to (currently) unavailable literature

This reference has been deleted.

P3, l9: there are a number of other relevant model description papers for FAMOUS - the original Jones et al ’05 paper should be additionally cited here at a minimum

This reference and Smith 2012 have been added.

p3,l14: Space needed between "Gordon et al.(2000)" and "HadCM3" p3,l17: "this in relatively poorly documented" - perhaps "is"?

Thank you for picking this up, it has now been correct.

P4, l7: since a point is made of the computational speed of HadCM3 et al over more modern models, some detail of the computational throughput/resource requirements of the different configurations would be useful, either here or elsewhere

This is a good point raised by both reviewers. We have included a new table of run speeds (table 3).

p4,l16,117: "L" only refers to the ocean resolution, and "H" only to the atmosphere

This is correct and has been clarified in the text.

p6,l15;p7,l29: WM - as a regular lat/lon grid is used in both components, Fourier filtering of higher wave-number dynamics is done in both models in certain latitude ranges.

A sentence has been added to each of these sections (atmosphere and ocean) to state this.

p8,l6/p8,l25: WM - the rigid lid formulation requires the pre-specification of "islands" around which the barotropic circulation may occur. The standard MetOffice HadCM3, at least, does not allow mass transport through the Bering St because of this, which may affect the AMOC stability characteristics.

The formulation of islands in the model is discussed in Section 4.1.4. Here, we add: “The barotropic solver requires the pre-specification of "islands" around which the barotropic circulation may occur (See Section 4.1.4)”

(Some of our current work on development is allowing mass transport through the Bering Strait and it does indeed have a big effect on the AMOC).

p8,l30: WM - the virtual salinity flux is calculated using a globally constant reference salinity, which can distort the local response to the surface water forcing

Thank you for pointing this out, we’ve added this sentence to the paper.

p9,l30: despite what many generations of UM code and documentation has asserted, the soil hydrology in all versions of MOSES is apparently not derived from Clapp and Hornberger '78, but Brooks and Corey '64 (see eg footnote at http://juleslsm.github.io/vn4.2/namelists/jules_soil.nml.html). The model still names everything with Clapp-Hornberger, so it would seem unhelpful to readers to start referring to Brooks and Corey when they won't find these names in the model, but this might be an opportunity to stop propagating the C-H misinformation.

Interesting. I was not aware of this history. A quick reading of the relevant papers suggests that Brooks and Corey devised the expression but C-H calculated more values. We have added a sentence and reference to Brooks and Corey.

p10,l4: I recall presentations by Valdes some years ago that appeared to show statistically different climates from the "same" HadCM3 ported to different platforms. Am I misremebering, or has this issue been cleared up to the authors' satisfaction?

Well remembered. This problem has indeed been sorted. There were two problems. One was the compiler bug (mentioned above) and the second was related to a rather strange issue related to reconfiguration. We now get the same climate (statistically) for HadCM3-M1 and our PMIP2 simulations are the same too.

p11,l4: MOSES2.2 can be used in FAMOUS, although most published FAMOUS papers use the MOSES1 variant.

We've added a comment about this to the text.

p11,l21: WM - MOSES2.2 has two modes of operation. One functions as described here (calculating the exchange for each tile, then averaging the fluxes), and the other aggregates the different tile properties together /before/ doing one calculation of the average flux (see Essery et al 03). I assume the authors use the first mode - FAMOUS (eg the Williams et al '13 they cite) uses the second mode

This is correct, we use the first mode. Essery et al. 2003 note that there are negligible differences in the result from the two approaches but that conceptually the first mode is easier within a GCM.

p21-: from this point, some model names have an N or H appended (eg HadCM3-M2.2N) - I don't see where this is explained

Corrected.

p23,l12: Beware FAMOUS-M2.2! I believe new issues have very recently been found with the long-term drift of the climate of the model described in Williams et al (Smith, pers.comm) that may play into this sort of bias and require significant retuning. Do you know what the AMOC is doing in that run?

We have performed a 1000 year run and the surface temperature and AMOC do not appear to have a drift (though the AMOC does have a lot of variability, see below) but we are aware of the issues that the reviewer alludes to. However, since these are still being investigated and FAMOUS is largely documented elsewhere, we have not amended the text.

p25,l8: The section title is a little misleading, since this section is purely about heat transports. On the subject of pure TOA fluxes: I believe HadCM3 is known to have compensating biases in TOA short- and longwave fluxes, linked to known problems in the clouds. WM?

This section has been renamed “Horizontal heat transports”.

p27,l9: At shorter timescales, however, FAMOUS was found to have high levels of variability in the AMOC when compared to the RAPID data (Balan Sarojini et al, Ocean Science 2011)

We have added this to the end of this sentence and cited Balan Sarojini et al.

p29,l17: the effectively shut Bering St may play a role here too

Pardaens et al (2003) ran a sensitivity study allowing flow equivalent to an open Bering strait and found that it had little impact on the salinity error but did have a big impact on the AMOC. We are actively working on improving the Bering St and, if satisfactory, will probably be an update to our version.

p29,l23-25: the end of this paragraph is phrased in rather too certain a manner concerning the reliability of the AMOC stability metrics presented eg in Liu et al’14 for my taste. It is too definite to claim that observations indicate that "the AMOC is in a bistable regime". Theoretical metrics have been derived that suggest that this *may* be the case, but they are some way from being proved definitive.

We have rephrased the sentence to lower the emphasis placed on the certainty of the “observations”

p33,l1: Reading this section, the uninitiated might expect that they would be able to obtain and run the useful models described here themselves. In fact, they would have great difficulty even viewing the effective model code, given the web of libraries, patches and options (all, technically, supplied here) that the UM is built from. That is not the fault of the present authors, and neither is the fact that this well-known model effectively has no support or distribution mechanism. But I think that some warning should be given that the copious information supplied in this section is *not* really sufficient to run the model oneself, and perhaps contact details (maybe for Bryan Lawrence, Director of Models and Data at NCAS) for someone to start with if a reader really did want to get help installing or running the system for their own use?

The text has been updated to make it clear that the code can only be viewed at the link given and to direct the reader to the UM Partnership Team for enquiries about using the model.

The version of HadCM3 used regularly within the group at Bristol has clearly branched from the Met. Office's original version. I think that this documentation (Valdes 2017) of it is a worthwhile contribution to GMD. I appreciate the provision of the source modifications and a list of the simulations. I would hope that in the final version there could be a link to the simulation output on the BRIDGE webpage. I had a couple queries, but suspect those could be addressed with revised sentences. I also found the figures and tables a little too small to be seen well on a printed version.

Thank you to Chris for his insightful comments and support of this paper.

2 Queries

Table 1 implies that two RHCrit values are due to level dependence rather than land/sea.

This is correct, the values of RHCrit in Table 1 do vary with level.

Table 2 shows that MOSES2.2 (and hence TRIFFID) either is or cannot be used with the atmosphere only model. I wasn't sure why.

We have changed the table to clarify that M2.2 *could* be used with HadAM3 and HadRM3, however it has not been done at Bristol (though would be simple to implement).

What is the 3A spectral scheme (p7, L19)

Spectral scheme 3A is a radiation scheme developed by the Met Office for version HadCM3 of the UM which is designed to treat both long- and shortwave schemes in a common, flexible framework as far as possible and is documented in Ingram, Woodward and Edwards (1997; Unified Model Documentation Paper No. 23: Radiation, http://cms.ncas.ac.uk/documents/vn4.5/p023.pdf). We have changed how we introduce this term and improved the referencing to make it clear where this term originates.

You state the soil layer thicknesses are a function of soil heat capacity and conductivity on p9, yet provide their depths on the p12. Are they constant across the globe?

The sentence written in the original manuscript has confused the reviewer: it is the amount of frozen soil moisture that is a function of soil heat capacity and conductivity, not the soil layer thickness. We have rewritten this sentence to make this clearer.

Are the surface types, LAI etc prescribed for each gridpoint (so 2D) or just each PFT.

The fractions of surface types and and values of LAI, canopy height and canopy conductance are specified at all grid points. The vegetation fractions and these other vegetation parameters will be updated by TRIFFID if used. Other parameters are hard-wired in the code. The text has been updated to clarify this.

You state moving from MOSES2.1 to MOSES2.2 have "particularly big" effect. Can you either quantify or refer to later section.

On reflection, we agree that ‘particularly big’ isn’t a helpful quantification, so we’ve changed this phrase to merely state that it affects the temperature.

In section 4.1.3 and the Table 2, you state that there is a ratio of solar radiation components that is 0. Is this correct, and what does it mean physically.

An explanation of this has been added to section 4.1.4 and table 1. It basically controls the way that the solar flux penetrates into the ocean and has been renamed to solar penetrative flux.

Section 4.1.5 seems redundant as only the sea-ice diffusivity in Table 1 hasn’t been explicitly mentioned.

We agree and have restructured the section and removed the mention of this.

What is the updating frequency of HadRM’s lateral boundary conditions - you state 6 hours on p17 and 3-4 hours on p18.

This can vary between 3 - 6 hours. In these simulations we used 4-hourly updating.

Please provide a little more information about the HadRM3 vs HadAM3 diffusion - there are 2 parameters in Table 1 and isn’t clear to me what they mean.

An explanation of these terms has been added to section 2.2.

Some Figures (e.g. Fig2) appear to use different acronyms to the text.

Thank you for pointing this out. The heading on Figure 2 has been updated to match the caption. The other figures have also been checked for consistency.

The discussion on p24 about leeward precipitation appears to be in the opposite sense as shown in figure 4.

We have clarified this section to highlight that we are discussing the bias in the gauge stations themselves rather than the models.

Section 5.1.3 does not discuss TOA flux - rather heat transports.

We have amended section title to “Heat transport”

Fig 5 contains no explanation of the gray lines.

We have added into the figure caption that these lines represent some of the CMIP5 models.

p27, l7. It isn’t clear to me what is meant by "larger annual variation" - within year or between years.

We have clarified these sentences as follows:
“HadCM3BL shows larger year to year variability than the observations: approximately twice as large as that in the observations. This results in years with a lower minimum volume transport than are seen in the observations. FAMOUS model variants tend to underestimate

the year to year variation by approximately (50%), although this is in contrast to the study of Sarojini 2011 who showed that FAMOUS exhibited greater short-term variability than the RAPID-MOCHA array. HadCM3B variants have a realistic year to year variability at least in the upper (1500 m) of the ocean.”

The references to Fig 9e/f need correcting
Apologies, these were left over from a previous version and have now been amended.

The discussion in Section 5.3.2 doesn’t necessarily recognise the big peak in the MODIS data at 25oN - rather it considers the feature symmetric around the Equator.

We did refer to the ‘subtropical spikes in productivity’, but we have rephrased this to emphasise this spike at ~20N.

3 Sentence Suggestions

p2, l2. I wonder if the final sentence of the abstract should use "predominantly"

rather than "particularly".

Changed.

p2, l9. Flux corrections are not often discussed anymore, so I think you need to

explain a little more

We have added a short sentence here to clarify the role of flux adjustments.

p7, l3. "which determine" occurs twice in quick succession.

This has been rephrased.

p7, l13. Be explicit that this relates to RHCrit in Table 1.

We have added RHCrit to the sentence.

p7, l26. you may want to consider giving the origin of some of the default fields.

We have expanded the text accordingly.

p8, l14-15. This feels like duplication of prior sentence.

Deleted.

p8, l24. "direct" -> "dynamic"?

Corrected.

p9, l10. The end of this sentence reads awkwardly.

This has now been modified.

p10, l6. Reference and explanation of the Visbeck scheme?

We added a reference a brief outline of the purpose of the 'Visbeck' scheme.

p10, l11. "numerous" -> "multiple" and incorporate ref to supplementary info.

Word changed and reference added to the supp. Information at the end of the list.

p10, l19. wrong section reference

This has been corrected.

p10, l24. This sentence has too many clauses. Split into two.

Done.

p11, l21-24. This reads awkwardly

Repharased.

p13, l16. Could this be related to section 3.1.2?

There is a partial link, which we have added to the text.

p19, l16. Rename section to "surface temperature patterns"?

Done.

p19, l32 " of 2" -> ", by 2"

Done.

Fig2 caption. section symbol §in a variable name

This has been corrected.

p25, l20. Spell FAMOUS correctly

Corrected.

p26, l2. You may want to add that this is despite the OHT biases.

We have added this statement to the end of the sentence.

p29, l15. remove paragraph break

Done.

p30, l7. I was unsure what "this version" of HadCM3 was from previous sentence. Refer to figure panel

We have amended the sentence to make reference to the HadCM3 models shown in Figure 9.

p30. l11. Why only this resolution?

The incorrect modelling of C4 grasses at the mouth of the Amazon by HadCM3 is related to the precipitation bias seen in Fig 4. The text has been updated to make this connection more clearly.

p31, l10-12. Can you refer back to the definition of these different terms?

Done.

p31, l24 "and which" -> ". We additionally show that"

Done.

The BRIDGE HadCM3 family of climate models: HadCM3@Bristol v1.0

Paul J. Valdes[1], Edward Armstrong[1], Marcus P. S. Badger[1,2], Catherine D. Bradshaw[3], Fran Bragg[1], Michel Crucifix[4], Taraka Davies-Barnard[1], Jonathan J. Day[5], Alex Farnsworth[1], Chris Gordon[6], Peter O. Hopcroft[1], Alan T. Kennedy[1], Natalie S. Lord[1], Dan J. Lunt[1], Alice Marzocchi[7], Louise M. Parry[1,8], Vicky Pope[10], William H. G. Roberts[1], Emma J. Stone[1], Gregory J. L. Tourte[1], and Jonny H. T. Williams[9]

[1]School of Geographical Sciences, University of Bristol, Bristol, UK
[2]School of Environment, Earth and Ecosystem Sciences, The Open University, Milton Keynes, UK
[3]Applied Science group, Met Office Hadley Centre, Exeter, UK
[4]Earth and Life Institute, Georges Lemaître Centre for Earth and Climate Research, Université catholique de Louvain, Louvain-la-Neuve, Belgium
[5]Department of Meteorology, University of Reading, Reading, UK
[6]Centre for Climate Research Singapore, Meteorological Service Singapore, Singapore
[7]Department of the Geophysical Sciences, The University of Chicago, Chicago, IL, USA
[8]Scottish Environment Protection Agency (SEPA), Perth, UK
[9]National Institute of Water and Atmospheric Research (NIWA), Wellington, New Zealand
[10]Met Office, Hadley Centre, Fitzroy Road, Exeter UK

Correspondence to: Paul J. Valdes (p.j.valdes@bristol.ac.uk)

Abstract.

Understanding natural and anthropogenic climate change processes involves using computational models that represent the main components of the Earth system: the atmosphere, ocean, sea-ice and land surface. These models have become increasingly computationally expensive as resolution is increased and more complex process representations are included. However, to gain robust insight into how climate may respond to a given forcing, and to meaningfully quantify the associated uncertainty, it is often required to use either or both of ensemble approaches and very long integrations. For this reason, more computationally efficient models can be very valuable tools. Here we provide a comprehensive overview of the suite of climate models based around the coupled general circulation model HadCM3. This model was ~~originally~~ developed at the UK Met Office and has been heavily used during the last 15 years for a range of future (and past) climate change studies but ~~is now largely being replaced by more recent~~ has now been largely superseded for many scientific studies by more recently developed models. However, it continues to be extensively used by various institutions, including the BRIDGE (Bristol Research Initiative for the Dynamic Global Environment) research group at the University of Bristol~~and elsewhere. Over time, adaptations have been made~~ , who have made modest adaptations to the base HadCM3 model over time. These adaptations mean that the original documentation is not entirely representative, and several other relatively undocumented configurations are in use~~which now differ from the originally described model versions~~. We therefore describe the key features of a number of configurations of the HadCM3 climate model family, ~~including the~~ which together make up HadCM3@Bristol version 1.0. In order to differentiate

variants that have undergone development at BRIDGE, we have introduced the letter B into the model nomenclature. We include descriptions of the atmosphere-only model (~~HadAM3~~HadAM3B), the coupled model with a low resolution ocean (~~HadCM3L~~HadCM3BL), the high resolution atmosphere only model (~~HadAM3H),~~ HadAM3BH), and the regional model (~~HadRM3)and a fast coupled model (FAMOUS), which together make up HadCM3@Bristol version 1.0~~HadRM3B). These also include three versions of the land surface scheme. By comparing with observational datasets, we show that these models produce a good representation of many aspects of the climate system, including the land and sea surface temperatures, precipitation, ocean circulation and vegetation. This evaluation, combined with the relatively fast computational speed (up to ~~2000×~~ 1000× faster than some CMIP6 models), motivates continued development and scientific use of the ~~HadCM3~~ HadCM3B family of coupled climate models, ~~particularly~~ predominantly for quantifying uncertainty and for long multi-millennial scale simulations.

1 Introduction

This paper describes the variants of the HadCM3 family of climate models ~~,~~ (all of which can be classed as General Circulation Models, GCMs), produced by the UK Hadley Centre/~~Meteorological~~ Met Office, and which remain in regular use by a number of research groups including the Bristol Research Initiative for the Dynamic Global Environment group (BRIDGE, http://www.bristol.ac.uk/geography/research/bridge). HadCM3 originated in the late 1990s with developments to the atmosphere model, HadAM3 (Pope et al., 2000)~~, followed by developments~~ . Together with improvements to the ocean ~~model (Gordon et al., 2000). The resulting coupled model,~~ GCM, this enabled the development of HadCM3 ~~,~~ (Gordon et al., 2000), which was one of the first ~~models~~ coupled atmosphere-ocean GCMs that did not require flux correction to maintain a reasonable present-day climate~~. It was~~ , i.e. the artificial adjustments of water, heat and momentum in order to maintain a stable climate. It has been extensively used for scientific studies of future climate change (e.g., Stott et al., 2000; Johns et al., 2003; Smith et al., 2007; Stott a is heavily cited, including in the 2007 IPCC report (Solomon et al., 2007) and was still included in the 2013 report~~, though~~. The family of models have the advantage of now being very well-known in terms of their strengths and weaknesses, as numerous studies have shown and classified model biases and forecast skill at representing the mean climate state as well as variability (e.g., Toniazzo et al., 2007; Spencer et al., 2007). The model family has now been superseded by the HadGEM2 (HadGEM2 Development Team, 2011) and HadGEM3 (Williams et al., 2015) families of models.

Compared to more recent models, HadCM3 is relatively low resolution but continues to perform reasonably well, at least with respect to its mean climate ~~(Flato et al., 2013)~~(Flato et al., 2013; Reichler and Kim, 2008). It also has the great benefit of computational speed, being more than ~~2000×~~ 1000× faster than some ~~versions of UKESM~~of the most recent and complex versions of the UK Met Office Unified Model (UM). This computational speed is particularly valuable for long-term simulations (necessary for many palaeoclimate simulations, studies which investigate the carbon cycle and the evolution of ice sheets) and for large ensembles (necessary for investigating the model's sensitivity to multiple parameters and quantifying the uncertainty in the model's response to forcing). Long model runs are also crucial for understanding unforced variability in the climate system (e.g., Collins et al., 2001).

Palaeoclimate simulations typically need many hundreds of model years to reach near-equilibrium in the surface and intermediate ocean and many thousands of years to reach equilibrium in the deep ocean. Moreover, there recently has been an increasing need to be able to consider the transient behaviour of past climate change. This has previously been tackled using the HadCM3 family of models by using either multiple "snapshot" simulations ~~(e.g., Singarayer and Valdes, 2010; Lunt et al., 2016; Marzocch~~ by performing fully transient simulations for ~~multi-millennial time scales (e.g., Hopcroft et al., 2014)~~multi-centennial or -millennial time scales (e.g., Tett et al., 2006; Hopcroft et al., 2014).

Faster models are also invaluable for investigating the sensitivity and robustness of results to changes in the initial ~~and boundary~~ conditions of the model ~~as~~ and changes in boundary conditions such as topography as numerous simulations can be performed ~~(Roberts and Valdes, 2017; Roberts et al., in review)~~(Roberts and Valdes, 2017). Additionally, they are ideal for ~~multiple runs to explore the sensitivity to a range of parameter values (e.g., using HadCM3 variants, Davies-Barnard et al., 2014; Armstrong~~ investigating anthropogenic changes on long timescales (Gregory et al., 2004; Ridley et al., 2005), and for performing perturbed parameter ensembles to rigorously calculate the probability density functions of either the mean or extreme climates ~~(e.g., using HadCM3 variants, Murphy et al., 2007; Booth et al., 2012; Jackson and Vellinga, 2013; Schaller et al., 2016)~~(e.g., using HadC Computational speed also aids more speculative studies. For instance, many early geoengineering simulations were run using variants of the HadCM3 family of models (such as Ridgwell et al., 2009; Singarayer et al., 2009; Lunt et al., 2008; Irvine et al., 2010).

~~A~~ In response to the need for fast models~~was the development of~~, Earth System Models of Intermediate Complexity (EMICs) (Claussen et al., 2002) have been developed. These models frequently achieve their speed by heavily parameterising the atmospheric response, even though atmospheric processes transport two-thirds of the total heat from equator to pole and play a vital role in the hydrological cycle. It is, therefore, also important to have a class of fast models that is equivalent to full atmosphere-ocean General Circulation Models (GCMs). Some EMICS do represent the dynamics of the atmosphere, for instance LOVECLIM (Goosse et al., 2010) uses a 3-level quasi-geostrophic atmosphere. Similarly, FAMOUS (part of the HadCM3 family) includes a full primitive equation atmosphere but at low resolution ~~(Smith et al., 2008)~~(Jones et al., 2005; Smith et al., 2008; Smith, 2012). Hence the division between EMICs and full complexity models is becoming increasingly blurred and we consider that the HadCM3 family provides a further bridge in the spectrum of models between intermediate complexity models and full complexity, state-of-the art models.

Since its introduction, HadCM3 (and related models) has undergone a ~~substantial~~ number of changes, bug fixes and adaptations~~, such that few of the versions of the model used now are truly identical to their original model description~~. The original model described in Pope et al. (2000) and Gordon et al. (2000)~~.~~, i.e. HadCM3 ~~(and family) is still extensively used by many groups~~with MOSES1, is still used but now many other versions exist. Some groups have largely stuck to the standard release of the model (e.g., Stainforth et al., 2005). Other groups have incorporated a variety of bug fixes and scientfic changes (in particular, many papers have used a revised land surface scheme, MOSES2.1 (e.g., Dolan et al., 2015) but this ~~in~~ is relatively poorly documented.

Therefore in this paper, we aim to rectify this for the wide range of HadCM3 variants currently in use within the BRIDGE modelling group. Our implementations of the models have diverged from other versions and so here we aim to provide clear documentation of our version of each model. In order to do this more clearly, we use the nomenclature HadCM3B, in order to differentiate model variants that have undergone development at Bristol to those originally developed at the Met Office. We have followed a specific modelling philosophy in which we attempt to minimise the differences between model configurations, particularly when changing resolution. For instance, previously published descriptions of HadRM3 (Jones et al., 1995) and HadAM3H (Hudson and Jones, 2002), (Arnell et al., 2003) use slightly improved physics to HadCM3 but we choose to keep the same physics (except for specific changes related to resolution).

We include detailed descriptions of each module of the models, differences between variants and comparison with observations across a range of metrics. This will increase transparency, traceability, and scientific openness. By detailing the changes and variations of these models, and providing an extensive comparison to observational data, we hope to show that these models remain useful tools for climate simulation and are suitable for further scientific use. Furthermore, we shall show that despite their relative simplicity, the models simulate the modern climate with comparable accuracy to many of the latest CMIP5 models.

To this end, we first describe the "base" model, ~~HadCM3-M1~~which we term HadCM3B-M1. This is essentially almost identical to that of Gordon et al. (2000) but with some minor modifications made by BRIDGE detailed in Sect. 2.1), to which all the other models will be compared (Sect. 2). ~~Then we~~ As such, because it is largely simply bug fixes, it could be argued that HadCM3B-M1 is not a different model to the original but we include it for completeness. We then subsequently discuss different land surface schemes (Sect. 3), followed by model variants with different ocean or atmospheric resolutions (Sect. 4). Finally, we evaluate the models' performance when compared to observations and CMIP5 models, to show that they recreate many key aspects of the climate system, and show which models are more suitable for certain applications (Sect. 5).

1.1 Overview of HadCM3@Bristol

The family of models has at its core HadCM3. From this core, variants are derived according to resolution, land-surface scheme, and components. ~~We choose to split the family~~ In order to distinguish variants that have undergone further development at Bristol to those originally developed at the Met office, we include the letter **B** for Bristol in the model acronym. As discussed in the text, the changes between the Bristol and Met Office variants are small in some cases, however we believe they warrant documentation to remove ambiguity. The model family is then split into groups: ~~HadCM3, HadCM3L (HadCM3~~ HadCM3B, HadCM3BL (HadCM3B but lower ocean resolution), ~~HadAM3 (HadCM3~~ HadAM3B (HadCM3B but atmosphere-only), ~~HadAM3H (HadAM3~~ HadAM3BH (HadAM3B but higher resolution), ~~HadRM3 (HadAM3~~ HadRM3B (HadAM3B but regional), and FAMOUS (HadCM3L but lower atmosphere resolution).

FAMOUS is a low resolution model derived from HadCM3, sharing much of the same physics, but with some numerical modifications suitable for the low resolution and which give quicker run times. It is well documented elsewhere (Jones, 2003; Smith et al., 2008) and will not be described again in detail here, although some comparisons with FAMOUS are included for completeness.

Run times for M1 model versions are compared in table 3 for typical configurations. This demonstrates the efficiency of FAMOUS-M1 at around several modelled centuries per day on just 8 cores, and the relatively high computational cost of the two high resolution model versions (HadAM3BH and HadRM3B). This compares with 1.87 model years per day on 1152 cores for the higher resolution version of HadGEM3-GC2 (Williams et al., 2015).

The nomenclature adopted for the HadCM3@Bristol model variants is Had⟨*Com*⟩~~M3~~M3B⟨*Res*⟩-⟨*Land*⟩⟨*Veg*⟩, where:

⟨***Com***⟩ (components) is one of:

- **A** – atmosphere-only model
- **C** – coupled model
- **R** – regional model

⟨***Res***⟩ (resolution) is one of:

- **L** – lower than standard resolution ocean
- **H** – higher than standard resolution atmosphere
- ***blank*** – standard resolution

⟨***Land***⟩ (land surface scheme) is one of:

- **M1** – MOSES1 land surface exchange scheme
- **M2.1** – MOSES2.1 land surface exchange scheme
- **M2.2** – MOSES2.2 land surface exchange scheme

⟨***Veg***⟩ (vegetation) is one of:

- ~~***blank***~~
- ***blank or N*** – no change to vegetation (i.e., static vegetation distribution)
- **E** – vegetation predicted using TRIFFID, but in “equilibrium” mode
- **D** – same as E above, but fully dynamic model

As such, the original “base” model described in Gordon et al. (2000) ~~, would be named HadCM3-M1 (although note that our version differs in several aspects from that in Gordon et al.,~~ which has undergone some minor modifications (see Sect. 2) is named HadCM3B-M1.

2 ~~HadCM3-M1~~HadCM3B-M1

This section describes the “core” model, ~~HadCM3-M1~~HadCM3B-M1, to which all other variants will be compared in this paper. This variant of the family was originally the most commonly used, and is still used for studies where the vegetation

Table 1. Summary of the key differences between model variants. For further details of these differences and description of the features common to all variants, see the relevant sections of the text. Note that ~~HadAM3~~ HadAM3B is identical to the atmosphere of ~~HadCM3~~HadCM3B.

Item	~~HadCM3~~HadCM3B	~~HadCM3L~~HadCM3BL	FAMOUS	~~HadAM3~~HadAMB3	~~HadAM3H~~HadAM3BH	~~HadRM3~~HadRM3B
Atmosphere						
Horizontal Resolution (n)	96×73	96×73	48×37	96×73	288×217	Varies with selected region
Horizontal Resolution (deg)	$3.75^\circ \times 2.5^\circ$	$3.75^\circ \times 2.5^\circ$	$7.5^\circ \times 5^\circ$	$3.75^\circ \times 2.5^\circ$	$1.25^\circ \times 0.83^\circ$	$0.4425^\circ \times 0.4425^\circ$ or $0.22^\circ \times 0.22^\circ$
Vertical Resolution	19 levels	19 levels	11 levels	19 levels	30 levels	19 levels
Timestep (mins)	30	30	60	30	10	5 or 2
Dynamics sweeps/physics timestep	1	1	1 or 2	1	2	1
Max wind test for half timestep dynamics ($m\,s^{-1}$)	—	—	—	—	240	—
Convective precipitation grid box fraction (conv_eps)	0.3	0.3	0.3	0.3	0.3	0.65 or 1.0
Large scale precipitation grid box fraction (ls_eps)	1.0	1.0	1.0	1.0	1.0	0.75 or 1.0
Boundary layer top and number of levels (eta/level)	0.835 / 5	0.835 / 5	0.9 / 3	0.835 / 5	0.8 / 6	0.835 / 5
Cloud levels (eta/level)	0.02 / 18	0.02 / 18	0.125 / 10	0.02 / 18	0.02 / 29	0.02 / 18
Pure pressure level start (eta/level)	0.04 / 17	0.04 / 17	0.06 / 11	0.04 / 17	0.04 / 28	0.04 / 17
Gravity wave drag start (eta/level)	0.956 / 3	0.956 / 3	0.9 / 3	0.956 / 3	0.956 / 3	0.956 / 3
Surface gravity wave constant (m)	2.0×10^4	2.0×10^4	2.0×10^4	2.0×10^4	1.6×10^4	2.0×10^4
Trapped lee wave constant ($m^{-3/2}$)	3.0×10^5	3.0×10^5	3.0×10^5	3.0×10^5	2.4×10^5	3.0×10^5
Filtering safety multiplying factor	0.01	0.01	0.011	0.01	0.1	—
Filtering wave numbers checked every	1 timestep	1 timestep	1 timestep	1 timestep	6 hours	—
Steep slope horizontal diffusion off until pressure level (kPa)	20	20	20	20	20	50
Diffusion coefficient ($m^6\,s^{-1}$) *	5.47×10^8	5.47×10^8	4.19×10^9	5.47×10^8	4.0×10^7	1.7×10^7
Diffusion power (dimensionless) *	~~3~~6	~~3~~6	~~4~~8	~~3~~6	~~2~~4	~~2~~4
Humidity diffusion coefficient ($m^6\,s^{-1}$ / $m^4\,s^{-1}$) *	5.47×10^8 1.5×10^8	5.47×10^8 1.5×10^8	2.4×10^8	5.47×10^8 1.5×10^8	2.0×10^7 4.0×10^7	1.7×10^7
Humidity diffusion power (dimensionless) *	~~3~~6 ~~2~~4	~~3~~6 ~~2~~4	~~2~~4	~~3~~6 ~~2~~4	~~2~~4	~~2~~4
RHcrit *	0.95 0.7	0.95 0.7	0.91 0.687	0.95 0.7	0.95 0.8	0.91 0.84 0.95
Ocean						
Horizontal Resolution (n)	288×144	96×73	96×73	—	—	—
Horizontal Resolution (deg)	$1.25^\circ \times 1.25^\circ$	$3.75^\circ \times 2.5^\circ$	$3.75^\circ \times 2.5^\circ$			
Vertical Resolution	20 levels to 5500 m	20 levels to 5500 m	20 levels to 5500 m	—	—	—
North Atlantic Bathymetry	Standard Met. Office	No Iceland	No Iceland	—	—	—
Vertical ~~Diffusion~~ Tracer Diffusivity	Richardson Number dependence	Constant background ~~mixing rate~~ value	Constant background ~~mixing rate~~ value	—	—	—
Coefficient for Solar Penetration (ratio)	0	3.8×10^{-1}	3.8×10^{-1}	—	—	—
Horizontal Momentum Diffusion Coefficient ($m^2\,s^{-1}$)	3×10^3	1.5×10^5	1.5×10^5	—	—	—
~~Horizontal Tracer Diffusion Coefficients~~ Isopycnal Diffusion Coefficients ($m^2\,s^{-1}$)	Visbeck et al. (1997) latitudinally varying scheme	Constant values	Constant values	—	—	—
Sea Ice Diffusion ($m^2\,s^{-1}$)	6.7×10^2	2.0×10^3	2.0×10^3	—	—	—

* Level dependent parameters (where multiple values are given, this indicates the range from surface to top-of-atmosphere (TOA))

is known, and as such can be prescribed, where relatively short simulations are sufficient for the science questions being addressed, and where the ocean plays a critical role and as such high resolution is desirable (e.g., Bragg et al., 2012).

Table 2. Availability of alternative land surface schemes.

Item	HadCM3 HadCM3B	HadCM3L HadCM3BL	FAMOUS	HadAM3 HadAM3B	HadAM3H HadAM3BH	HadRM3 HadRM3B
MOSES1	HadCM3-M1 HadCM3B-M1	HadCM3L-M1 HadCM3BL-M1	FAMOUS-M1	HadAM3-M1 HadAM3B-M1	HadAM3H-M1 HadAM3BH-M1	HadRM3-M1 HadRM3B-M1
MOSES2.1	HadCM3-M2 HadCM3B-M2.1	HadCM3L-M2 HadCM3BL-M2.1	—	HadAM3-M2 HadAM3B-M2.1	HadAM3H-M2 HadAM3BH-M2.1	HadRM3-M2.1 *
MOSES2.1 TRIFFID (D and E)	HadCM3-M2 HadCM3B-M2.1	HadCM3L-M2 HadCM3BL-M2.1	—	HadAM3-M2 HadAM3B-M2.1	HadAM3H-M2 HadAM3BH-M2.1	— *
MOSES2.2	HadCM3-M2 HadCM3B-M2.2	HadCM3L-M2 HadCM3BL-M2.2	FAMOUS-M2.2	— *	— *	— *
MOSES2.2 TRIFFID (D and E)	HadCM3-M2 HadCM3B-M2.2	HadCM3L-M2 HadCM3BL-M2.2	FAMOUS-M2.2	— *	— *	— *

* Variant currently does not exist but there is no barrier to creation

Table 3. Computational performance of M1 configurations.

	HadCM3B-M1	HadCM3BL-M1	FAMOUS-M1	HadAM3B-M1	HadAM3BH-M1	HadRM3B-M1
Cores	16	16	8	16	64	16
Speed (model years/day)	47	85	450	109	5	6
Cost (model years/day/core)	3.0	5.3	56.3	6.8	0.07	0.4

This model is a three dimensional, fully dynamic, coupled atmosphere-ocean global climate model without flux adjustment. Our version of the model is very similar to that described by Gordon et al. (2000). Our aim is to provide a brief description of this core model, and a full description of how our version differs from that in Gordon et al. (2000) followed by a brief description of the core model. A full description of the Gordon et al. (2000) version can be found in the UK Met Office technical notes http://cms.ncas.ac.uk/wiki/Docs/MetOfficeDocs; the base model code is currently available to view at http://cms.ncas.ac.uk/code_browsers/UM4.5/UMbrowser/index.html., but it should be noted that additional modifications are required for the full scientific definition of the model as described here (see Sect. 7).

2.1 Atmosphere (HadAM3)Modifications for the Bristol Version of HadCM3B-M1 and HadAM3B

We have benchmarked the standard version of HadCM3-M1 supplied by the UM Met office against existing model results from the published Hadley centre version of Gordon et al. (2000) and confirmed that we could reproduce the results within the normal statistical variability of the model. Subsequently a few relatively minor changes have been made. These include:

2.1.1 HadCM3B-M1

– Correction of a small bug in the Visbeck horizontal eddy mixing scheme (Visbeck et al., 1997) which was originally included in the standard configuration of the model to ensure compatibility with previous versions.

– Use of versions of the radiation and primary field advection schemes that are scientifically identical to the standard version and which make the model faster but are not bit reproducible.

2.1.2 HadAM3B

– Fixes to a few array bounds errors (which may or may not have an impact on the scientific results).

– Multiple other bug fixes which did not change the science but corrected problems with some aspects of the code and diagnostic outputs.

– There were two small bugs in the conservation of atmospheric mass, and the computation of vertical velocity, fixes to which are not included in the standard release version of HadAM3 but are included in HadCM3B. We include these bug-fixes in all versions of the code so that our atmosphere model (HadAM3B) is 100 % identical to the atmosphere component of our version of HadCM3B.

– There is also another important code fix (Steenman-Clark, pers. comm) which is vital to include. If this is not included, then some compilers will lead to a large (e.g., 0.75 °C bias in global mean surface air temperature) error in mean climate.

These modifications are included in the supplementary information. The overall impact of these changes on the climate simulation is very small.

2.2 Atmosphere (HadAM3B)

The atmosphere component of ~~HadCM3~~ HadCM3B is almost identical to the atmosphere component of ~~HadAM3~~HadAM3B, which is the atmosphere-only variant with fixed sea surface temperatures (SST). ~~HadAM3~~ The modifications made at Bristol are highlighted in Sect. 2.1, beyond which the model is the same as that described by Pope et al. (2000). HadAM3B has a Cartesian grid with a horizontal resolution of 96×73 grid points (3.75° longitude×2.5° latitude) with 19 hybrid levels (sigma levels near the surface, changing smoothly to pressure levels near the top of the atmosphere) in the vertical (Simmons and Strüfing, 1983) and uses a 30-minute timestep. ~~HadAM3~~ HadAM3B solves the primitive equation set of White and Bromley (1995) which includes certain terms necessary to conserve both energy and angular momentum. Equations are solved through the use of a grid-point scheme, specifically the Arakawa staggered B-grid (Arakawa and Lamb, 1977), on a regular latitude-longitude grid in the horizontal. At high latitudes, Fourier filtering of higher wave-number dynamics is used to remove subgrid-scale variability. A split-explicit time scheme conserves mass, mass weighted potential temperature, moisture and angular momentum, and ensures the reliability for solving equations on long time scales, which is particularly important for climate modelling (Van der Wal, 1998).

As with any climate model, a number of parameterisation schemes are needed within ~~HadAM3~~ HadAM3B to represent certain physical processes which occur on sub-grid scales:

– Precipitation is dealt with in two schemes: i) the large-scale precipitation scheme, and ii) the convection scheme. The large-scale precipitation scheme removes cloud water resolved on the grid-scale, i.e., frontal precipitation. This is done via a simple bulk parameterisation scheme converting water content into precipitation (Wilson, 1998). The convection scheme (Gregory et al., 1997) uses a mass-flux scheme with the addition of convective downdrafts.

– A first order scheme for turbulent vertical mixing of momentum and thermodynamic quantities is used within the boundary layer, which can occupy up to the first five layers of the model. Sub-gridscale gravity wave and orographic drag parameterisations include the impact of orographic variance anisotropy (Gregory et al., 1998). The scheme comprises four elements: i) "Triggering" which determines whether the physical conditions within the grid-box constitute convection taking place; ii) "Cloudbase closure" ~~which determines~~ controlling the intensity of convection which is determined by the mass transported through the cloudbase; iii) a transport model where temperature, moisture, wind fields and thus precipitation, are determined and iv) "Convective cloud scheme" where cloud fractions derived from convection are calculated which will be used by the radiation scheme (Grant, 1998).

– In the real world, clouds are formed on scales far below that of the coarse grid used in ~~HadAM3~~HadAM3B, therefore there is the need for a statistical parameterisation of this variable. Probability Density Functions are used on the total water content over the grid-box mean to parameterise cloud amount/distribution and longevity (Bushell, 1998). Clouds are modelled as either water, ice or mixed-phase when the temperature in the model level is between 0 °C and −9 °C. Clouds form when the mean plus the standard deviation of the grid-cell moisture content exceeds a threshold of relative humidity (see RHCrit in Table 1 for numerical values). This cloud water content can then be used to produce a cloud fraction for each grid-box (Bushell, 1998). The threshold of total water content for precipitation to occur varies between land and ocean cells to account for the different levels of available cloud condensation nuclei. The scheme uses temperature through the vertical levels to determine the ice and water phases to determine cloud water content.

– Radiation is represented using the radiation scheme of Edwards and Slingo (1996). This scheme has six ~~short-wave and eight long-wave~~ short wave and eight long wave bands and represents the effects of water vapour, carbon dioxide, ozone and minor trace gases. A background aerosol climatology following Cusack et al. (1998) increases the atmospheric absorption of ~~short-wave~~ short wave radiation relative to previous versions representing a significant improvement. The ~~3A longwave and shortwave spectral scheme is used and~~ long wave and short wave spectral scheme used ("3A" of Ingram et al., 1997) is an improvement over the previous versions as it allows the freedom of choices of cloud parameterisation, gases and aerosols to be included through spectral input files (Edwards, 1998).

– Horizontal diffusion takes the form $k\nabla^N$ where both k and N can vary with vertical levels and with variable. The standard resolution of the model uses a formulation $k_1\nabla^6$ where $k_1 = 5.47\times10^8\ \mathrm{m^6\,s^{-1}}$ corresponding to a e-folding time scale for the two-grid wave of approximately 12 hours. The top most level in the model uses a stronger diffusion of the form $k_2\nabla^2$ where $k_2 = 4.0\times10^6\ \mathrm{m^2\,s^{-1}}$. Moisture also has stronger diffusion in the five levels below the top (approximately from 150 hPa) corresponding to $k_m \nabla^4$ where $k_m = 1.5 \times 10^8$ m^4 s^{-1}. The functional form and strength of diffusion for other resolutions is summarised in Table 1.

Boundary conditions for the model include the land sea mask, orography and its subgrid scale variability (originally derived from the US Navy updates 10' dataset), and a range of soil and vegetation parameters (originally derived from data in Wilson and Henderson-Sellers (1985). The model also needs to be initialised with soil moisture and snow cover (based on Willmott et al., 1985), and deep soil temperatures (empirically derived using Warrilow et al. (1986). When the model is run in atmosphere-only mode, i.e., HadAM3HadAM3B, sea surface temperature and sea ice (concentration and depth) are required to be prescribed. These can be derived from observational data or from coupled model simulations.

2.3 Ocean

The ocean component has a horizontal resolution of 288×144 grid points (1.25°×1.25°) (Gordon et al., 2000) . Therefore and, as with the atmosphere, also uses Fourier filtering at high latitudes. The higher resolution means that six ocean grid cells correspond to each atmosphere grid cell. In order to simplify the coupling of the atmosphere and ocean models, the land-sea mask is defined at the atmosphere resolution; therefore, the ocean model's coastlines appear relatively coarse. In the vertical there are 20 depth levels with finer definition at the ocean surface, with the top-most model layer being 10 m thick and the bottom-most 616 m thick. The ocean timestep is one hour. The ocean and atmosphere modules are coupled once a day with no flux adjustment necessary.

The ocean model is based on the model of Cox (1984) and is a full primitive equation, three dimensional model of the ocean. A second order numerical scheme is used along with centred advection to remove nonlinear instabilities. The Arakawa B-grid is used for staggering of tracer and velocity variables, allowing for more accurate numerical calculations of geostrophically balanced motion. It uses a rigid lid which eliminates fast external mode gravity waves found in the real ocean, thus allowing for longer timesteps, and with the result that there is no variation in the volume of the ocean. The barotropic solver requires the pre-specification of "islands" around which the barotropic circulation may occur (see Sect. 4.1.5).

As with the atmosphere, the ocean model also requires a number of parameterisations:

- The ocean mixed layer is represented by the Kraus and Turner (1967) model which assigns 15 % of gravitational potential energy and 70 % of wind-stress energy to turbulent kinetic energy, which is mixed out exponentially with depth. At all depths, five iterations of convective mixing are carried out at each timestep. Tracer and momentum mixing is modelled using the K-Theory scheme. Within the mixed layer a simplified version of the Large et al. (1994) scheme is employed: below this the Pacanowski and Philander (1981) K-Theory parameterisation is used.

- Further vertical mixing is provided at all depths as the sum of depth-dependent constant background term and a term dependent on Richardson number (Pacanowski and Philander, 1981).

– Momentum mixing is approximated using diffusion that is governed by a coefficient that consists of two terms: a constant background value and a term dependent on the local Richardson number. For tracers, diffusion increases with depth as detailed in Table A of Gordon et al. (2000).

– Horizontal eddy mixing of tracers is carried out using the isopycnal parameterisation of Gent and McWilliams (1990), with thickness diffusion coefficients modified following the method of Visbeck et al. (1997). Isopycnal mixing uses the Griffies et al. (1998) implementation of the Griffies et al. (1982) scheme. The along-isopycnal diffusion coefficient is 1000 $m^2 s^{-1}$. Horizontal mixing of momentum is performed using a latitudinally varying formulation which, coupled with the finer resolution of the ocean grid, enables western boundary currents to be resolved.

– There is no ~~direct~~ dynamic connection between the Mediterranean Sea and Atlantic Ocean so it is modelled as a “diffusive pipe” by completely mixing the easternmost point of the Atlantic with the westernmost point of the Mediterranean. Mixing occurs over the top 13 layers, to a depth of 1200 m, on the assumption that Mediterranean water will sink to at least this depth. A similar parameterisation is applied in the outflow of the Hudson Bay.

– Ice sheets are not modelled dynamically, therefore the snow accumulation on surface land ice points and over isolated water bodies must be balanced by loss through a notional iceberg calving that is represented as a time-invariant freshwater flux (which, because of the rigid lid, is converted to a virtual salinity flux). This is distributed around the edge of the ice sheets and polar oceans. The virtual salinity flux is calculated using a globally constant reference salinity, which can distort the local response to the surface water forcing. River runoff is instantaneously transferred to the ocean using a prescribed runoff map.

The modern bathymetry for the model is derived from the ETOPO5 reconstruction (Edwards, 1989) using a simple smoothing algorithm. The geometry of some significant channels is modified from the resulting coarse interpolation to ensure a more realistic model performance (Gordon et al., 2000). For example, the Greenland–Scotland ridge and Denmark Strait have significant sub-gridscale channels which are lost in the smoothing and so have been re-created by deepening channels (single cell width~~channels~~) in three locations along the ridge to reproduce the mean outflow to match observations, and the bathymetry around Indonesia is modified to ensure that flow occurs between Indonesia and Papua New Guinea but not between Indonesia and the mainland of Asia.

2.4 Sea ice

Sea ice is calculated as a zero layer model on top of the ocean grid. Partial cell coverage of sea ice is possible up to 0.995 in the Arctic and 0.98 in the Antarctic~~according to~~. This is based on the parameterisation of sea ice concentration ~~due to~~ from Hibler (1979). Ice forms primarily by freezing in leads, although ice can also form from snow falling on existing ice. It is assumed to freeze at the base of the sea ice at the freezing point of -1.8 °C. A constant salinity is assumed for ice, with the excess salt on melting/formation added as a flux into the ocean. Sea ice dynamics are simply parameterised: the surface wind stress over sea ice is applied to the ocean beneath the ice, and the ice thickness, concentration and accumulated snow then drift following the ocean currents in the top model layer (Gordon et al., 2000). The maximum depth that sea ice can reach due to convergence from drift is limited to 4 m in depth, although it may subsequently thicken further due to freezing. The albedo of sea ice is set at 0.8 for temperatures below -10 °C and 0.5 for temperatures above 0 °C, with a linear variation between these values.

2.5 Land Surface Scheme: MOSES1

The land surface scheme MOSES (Met Office Surface Exchange Scheme) is built upon the previous Met Office land surface scheme (UKMO) (Warrilow and Buckley, 1989). In the Gordon et al. (2000) version of HadCM3, MOSES Version 1 MOSES1, is used. A technical overview of MOSES1, a comparison to its predecessor (UKMO) and its climatological impact is provided by Cox et al. (1999).

In addition to calculating the fluxes of water and energy, MOSES1 incorporates the physiological impact of atmospheric carbon dioxide, water vapour and temperature on photosynthesis and stomatal conductance. It accounts for the effects of freezing and melting of soil moisture in four soil layers~~whose thicknesses are~~ , the proportion of frozen soil moisture being a function of the soil heat capacity and conductivity of the grid cell. Both vegetated and non-vegetated land surface types are characterised by a set of surface properties that are not updated during the model run. The canopy scheme is based on that used in Warrilow et al. (1986).

MOSES1 has two sets of prescribed land surface property attributes, which are input into the model via two external files. The soil attributes are volumetric soil moisture concentration at the wilting point, critical point, field capacity, and saturation, the saturated hydrological soil conductivity, the Clapp–Hornberger B exponent, the thermal capacity of soil, thermal conductivity of soil, and the saturated soil water suction. ~~The~~ (The Clapp–Hornberger exponent is a measure of the pore volume distribution and the formulation was originally devised in Brooks and Corey (1964)).The vegetation attributes are root depth, snow free albedo, stomatal resistance to evaporation, surface roughness, canopy water capacity, infiltration enhancement rate, deep snow albedo, leaf area index and canopy height of vegetation fraction. All of these attributes are derived from the Wilson and Henderson-Sellers (1985) data set.

2.6 ~~Modifications for the Bristol Version of HadCM3-M1~~

~~We have benchmarked the standard version of HadCM3-M1 supplied by the UM Meteorological office against existing model results from the published Hadley centre version of Gordon et al. (2000) and confirmed that we could reproduce the results within the normal statistical variability of the model. Subsequently a few relatively minor changes have been made. These include:~~

- ~~Correction of a small bug in the Visbeck scheme which was originally included in the standard configuration of the model to ensure compatibility with previous versions.~~

- ~~Use of versions of the radiation and primary field advection schemes that are scientifically identical to the standard version and which make the model faster but are not bit reproducible.~~

- ~~Fixes to a few array bounds errors (which may or may not have an impact on the scientific results).~~

- ~~Numerous other bug fixes which did not change the science but corrected problems with some aspects of the code and diagnostic outputs.~~

- ~~There were two small bugs in the conservation of atmospheric mass, and the computation of vertical velocity, fixes to which are not included in the standard release version of HadAM3 but are included in HadCM3. We include these bug-fixes in all versions of the code so that our atmosphere model (HadAM3) is 100 % identical to the atmosphere component of our version of HadCM3.~~

~~The overall impact of these changes on the climate simulation is very small.~~

3 Alternative Land Surface Schemes

Section ~~3~~ 2.5 describes the MOSES1 land-surface scheme which is used in the standard version of HadCM3. Here we describe two other versions, MOSES2.1 and MOSES2.2, as well as the vegetation component TRIFFID.

3.1 MOSES2

MOSES1 requires maps of vegetation properties, such as root depth and leaf area index~~which have~~ , to be prescribed (normally in a set of external files)~~, and as such~~. As such, it is not very suitable for an interactive vegetation model. As part of the process of developing a dynamic vegetation module for HadCM3, an upgraded land surface scheme, MOSES2, was also developed~~, called MOSES2~~. The first version of this scheme, MOSES2.1, is the original scheme used in early work with dynamic vegetation (Cox et al., 2000). This version was originally coupled to HadCM3LC (Cox et al., 2000), which is a flux-corrected low resolution version of HadCM3 which includes a carbon cycle. MOSES2.1 was further developed ~~by Michel Crucifix~~ for use in HadCM3 as part of the Paleoclimate Modelling Intercomparison Project Phase II (PMIP2) (Braconnot et al., 2007). Subsequently, a second version of MOSES2 was developed, MOSES2.2 (Essery et al., 2001, 2003) which was similar scientifically to MOSES2.1 but had improved code structure~~and~~ . This has become the initial core of the land surface model JULES (Best et al., 2011; Clark et al., 2011). At the University of Bristol, we have mainly used MOSES2.1, with MOSES2.2 only being used in a few specific contexts such as for investigating changes in atmospheric chemistry (Valdes et al., 2005; Beerling et al., 2011) because it can include additional parameterisations of isoprene emissions. MOSES2.2 ~~is also~~ can also be used in FAMOUS (Williams et al., 2013) though the majority of FAMOUS publications have used MOSES1.

A detailed discussion of the upgrades between MOSES1 and MOSES2.2 is provided in Essery et al. (2003) and a full and complete technical overview of MOSES2.2 in Essery et al. (2001). But so far there have been no clear comparisons as to how MOSES2.2 differs scientifically or technically from MOSES2.1, despite MOSES2.1 being the core version used at Bristol. The following sections aim to rectify this and clarify the differences between MOSES2.2 and MOSES2.1~~, after first outlining~~ . First we outline how MOSES2.2 differs from MOSES1.

3.1.1 Differences between MOSES2.2 and MOSES1

Compared to MOSES1, MOSES2.2 has major upgrades to all aspects of the land surface exchange and the surface radiation scheme Essery et al. (2003). The surface radiation scheme has an updated coupling between the land surface and atmosphere, including the calculation of surface net radiation and surface heat and moisture fluxes. MOSES2.2 allows fractional coverage of different surface types on a sub-grid scale. There are nine land-surface types explicitly modelled at a sub-grid scale, each with a set of characteristic parameters. MOSES2.2 can be fully coupled to the dynamic vegetation model TRIFFID (see Sect. 3.2) . The fractional coverage includes via the five plant functional types (PFTs): broadleaf trees, needleleaf trees, shrubs, C3 (temperate) grasses and C4 (tropical) grasses. Each have different values for a range of phenological characteristics including Leaf Area Index (LAI), albedo, and canopy interception. The remaining four are non-vegetated surface types; urban, inland water, bare soil, and ice. Excluding ice type, each land-surface gridbox grid box can be made up of any mixture of the other eight surface types. Land ice has to must have fractional cover of 0 or 1 only. The fractional coverage for each surface type is specified for each grid point from an external file. In addition, another file is supplied specifying the necessary parameters for the five vegetation types at each gridpoint: leaf area index (LAI), canopy height and canopy conductance (not PFT dependent). The vegetation fractions and parameters will be updated by TRIFFID if it is being used. Other PFT dependent parameters, including root depth and values of albedo under a variety of conditions, are hard-wired into the code.

Unlike In MOSES1, the surface energy balance in and moisture fluxes are calculated based on grid box average values of parameters (such as roughness length etc.). In MOSES2.2, the surface energy balance is explicitly solved for each surface type(tiles), and then weighted by the fractional area of the surface types within the grid box. This produces the gridbox grid box average surface temperature and soil moisture and fluxes of longwave, shortwavelong wave, short wave, sensible, latent and ground heat. In contrastAbove the surface, air temperature, humidity and wind speed on atmospheric levels above the surface are treated as homogeneous across the gridbox. Likewisegrid box. Similarly, soil temperatures and moisture contents below the surface are also treated as homogeneous. The aerodynamic surface roughness lengths are calculated explicitly according to the canopy height and the rate of change of roughness length with canopy height for each tile. This roughness length is used to calculate surface-atmosphere fluxes of heat, water, momentum and CO_2. The surface albedo determines the amount of downward shortwave short wave heat flux that is reflected at the surface. The surface albedo for fractional covered vegetated surface types (unweighted) is described by the snow-free and cold deep snow albedos. The soil albedo is defined according to colour and moisture content. LAI is also used in determining the surface albedo for surfaces covered by vegetation.

The hydrological cycle in MOSES2.2 is more similar to MOSES1 with small changes for the interactions with vegetation. However, it continues to treat each tile separately so extraction of water from the soil is calculated for each tile and then weighted summed to give the grid box average. Precipitation is partitioned into interception (via the canopy), throughfall, run-off and infiltration into the ground. Different parameters apply to each vegetation type. Canopy water refers to the precipitation intercepted by plant leaves available for free evaporation. MOSES 2 uses the same four soil layers as MOSES1, with thicknesses from the surface downwards set to 0.1 m, 0.25 m, 0.65 m and 2 m. Moisture content of the upper soil layer (0.1 m) is increased via snow melt and throughfall and decreased according to evaporation from the soil layer, flow of water into lower layers and draw up of water via plant roots. The extraction of water from any particular soil layer is proportional to the water lost by evapotranspiration reflecting the vertical distribution of roots. The five PFTs have different root depths, such that trees are able to access moisture from soil layers at deeper depths compared with grasses and shrubs. The soil moisture content and soil water phase changes and the associated latent heat describe the thermal characteristics of soil that determine, via discretised form of the heat diffusion equation, the subsurface temperatures. Subsurface soil temperatures are determined by the diffusive heat fluxes into and out of a soil layer and the heat flux advected from the layer by the moisture flux.

MOSES2 requires similar soil parameter inputs to MOSES1 although it additionally requires bare soil albedo and soil carbon content of the soils. However, the vegetation properties are very different. MOSES1 required inputs of grid box average LAI, root depth, etc., whereas MOSES2 requires prescribed inputs of the fractional types of each surface type, the LAI and canopy height of each vegetated PFT, and the overall canopy conductance. It also includes a disturbance fraction that represents agriculture. If using dynamic vegetation (TRIFFID), then these fields (except for disturbance) are only used for initialisation and the model will dynamically update them.

3.1.2 Differences between MOSES2.2 and MOSES2.1

There are a number of key differences between MOSES2.1 and MOSES2.2, and a number of smaller modifications between the versions. These major changes include:

- MOSES2.2 uses a spectral albedo scheme to calculate separately the diffuse and direct-beam surface albedos. This scheme is not used in MOSES2.1, although modifications can be added to include it.
- MOSES2.2 uses a spectral snow albedo model that includes a prognostic grain size that characterises the ageing of snow and its impact on snow albedo. This is not present in MOSES2.1.
- MOSES2.2 also introduces a new calculation of evapotranspiration from soil moisture stores, as well as a different parameterisation of bare soil evaporation
- Supersaturation in the soil layer is treated differently in the two versions of MOSES2. In MOSES2.2, supersaturation results in an increase in surface run-off. In contrast, supersaturation in MOSES1 and MOSES2.1 is managed via an increase in downward flow into the deeper soil layers and so is removed via subsurface runoff.

Tests carried out in which MOSES2.1 is gradually changed to MOSES2.2 show that the first two changes ~~have a particularly big effect on~~ affect surface temperature whereas the third difference substantially alters soil moisture. Supersaturation changes ~~have a big impact on~~ impact the partitioning of runoff between surface and sub-surface and also ~~have a small impact on~~ influence the soil moisture, to a lesser extent to the evapotranspiration changes.

There are also a number of smaller changes (such as using an implicit soil moisture scheme in MOSES2.2 compared to an explicit scheme in MOSES2.1 ~~(~~and MOSES1) but these do not result in a major change to the climate. MOSES2.2 also had some major restructuring of the Fortran code.

Additionally, in the default version of MOSES2.1 (used until recently), the rate of respiration increases almost exponentially with temperature (Tindall, pers.comm.). As a result, in some conditions such as during the Amazon dry season, respiration excessively increases and this decreases soil moisture which consequently inhibits tree growth. In MOSES2.2, the impact of temperature on respiration rate declines at high temperatures. This revised respiration rate reduces drying and dieback of trees. This has now become the default for the Bristol variant of ~~HadCM3-M2~~HadCM3B-M2.1 too.

3.2 TRIFFID

MOSES2.1 and MOSES2.2 both have the capacity to be run in coupled mode with a dynamic vegetation and terrestrial carbon cycle scheme, TRIFFID (Top-down Representation of Interactive Foliage and Flora Including Dynamics) (Cox et al., 1998; Cox, 2001). TRIFFID predicts the distribution and properties of global vegetation based on plant functional types using a competitive, hierarchical formulation. The performance and sensitivity of TRIFFID has been compared with a variety of other dynamic vegetation models (Sitch et al., 2008) and an updated version of TRIFFID is used in both the latest Coupled Model Intercomparison Project (CMIP5) model HadGEM2-ES (Collins et al., 2011) and in JULES (Clark et al., 2011).

In the model configurations presented here, TRIFFID is normally only used with MOSES2.1 because of a dry bias in MOSES2.2 which is manifested by an overly dry surface climate over the Eurasian continent in summer. This results in loss of vegetation if used with dynamic vegetation. The cause of this drying is unclear but is partially linked to the changes in evaporation and evaptranspiration parameterisations discussed above.

TRIFFID updates the five vegetation PFTs and the bare soil fraction, all of which can change dynamically. TRIFFID can be run in two different modes:

- Equilibrium mode, where TRIFFID runs for 50 years of TRIFFID for each 5 years of the climate model run. The fluxes between the land and the atmosphere are calculated and averaged over 5 years. This is particularly valuable for quick spin-up of the vegetation and soil carbon.

- Dynamic mode, where TRIFFID is run every 10 days. Fluxes are averaged over 10 days; as such high frequency variability is accounted for. This mode is the standard for full runs of the coupled model.

MOSES2 passes the averaged fluxes of carbon to TRIFFID which calculates the growth and expansion of the existing vegetation, and updates the land surface parameters based on the new vegetation distribution and structure. TRIFFID calculates areal coverage, leaf area index (LAI) and canopy height for five defined plant functional types (PFTs): broadleaf tree, needleleaf tree, C3 grass, C4 grass and shrub. These PFTs respond differently to climate and CO_2 forcing (e.g., C3 and C4 grasses use different photosynthetic pathways), and also impact differently on the physical properties of the land surface, i.e., possessing different aerodynamic roughness lengths and albedo properties. Broad and Needleleaf trees and C3 and C4 grasses react independently within the model due to their unique parameter sets. C4 plants use water more efficiently than C3 plants, requiring less water to produce the same amount of biomass. Overall, C4 plants have the highest critical humidity deficit and temperature range, meaning that in high temperature, low moisture environments they will do better than other PFTs, even though the competition model would normally favour trees.

All PFTs can co-exist within the same grid box, each possessing a fractional coverage that is equivalent to the population size. The fractional coverage co-existence approach allows smooth transitions to occur when the vegetation distribution changes rather than the sudden discontinuities that would occur in a "dominant" PFT only approach (Svirezhev, 2000). However, the Lotka–Volterra equations used in TRIFFID mean that each grid cell in the model tends to converge on one dominant plant functional type (Hughes et al., 2006). Competition is essentially based on a height hierarchy of trees > shrubs > grasses. Each terrestrial grid square has a small minimum content of each plant functional type, regardless of location and competition, as a "seeding" fraction (Cox, 2001). This ensures that no PFT can become extinct and can regenerate when conditions become appropriate. TRIFFID can specify areas of agricultural crops as C3 and C4 grasses, without competing land types (Cox, 2001).

The terrestrial Net Primary Productivity (NPP) is calculated by a coupled photosynthesis-stomatal conductance model (Cox et al., 1998). Factors affecting the rate of photosynthesis are the humidity deficit, the photochemically active radiation, soil moisture and LAI. The maximum rate of photosynthesis is directly related to the leaf temperature and the upper and lower temperatures for photosynthesis (defined individually for each PFT). Carbon is stored in the vegetation and soil stores.

The predicted vegetation in each grid box feeds back into the climate system in a number of ways, principally through evapotranspiration from the canopy, alteration of surface albedo, and through alteration of mixing at the boundary layer between the surface and the atmosphere (due to changes in roughness length).

4 Variants with Differing Resolution

4.1 ~~HadCM3L~~ HadCM3BL

~~HadCM3L~~ HadCM3BL comprises the same model components as ~~HadCM3~~HadCM3B, but with a lower resolution ocean which matches the standard atmosphere resolution of 96×73 grid points (3.75°×2.5°) ~~.~~ (Cox et al., 2000). Note that the Bristol version, HadCM3BL, is very different from the Met Office version. The Met Office version was mainly used for the early carbon cycle work (Cox et al., 2000) but required significant flux corrections to ensure that the Atlantic surface climate was reasonable. Our version does not require flux correction because of changes in bathymetry described below. It can be run with all versions of MOSES, with or without TRIFFID, in the same manner as ~~HadCM3~~HadCM3B. We tend to use ~~HadCM3L~~ HadCM3BL when long simulations are required. For instance, when the land-sea mask and/or bathymetry are substantially changed from those of modern, it can take many thousand years of ~~intergration~~ integration to get the deep ocean into equilibrium. As such, ~~HadCM3L~~ HadCM3BL has been used extensively for our pre-Quaternary climate modelling work (e.g., Marzocchi et al., 2015b; Bradshaw et al., 2015; Kennedy et al., 2015; Loptson et al., 2014).

The implementation of the atmosphere and land surface schemes is identical to ~~HadCM3~~HadCM3B. There are some differences in the ocean due to its lower resolution, some of which are substantive differences required either to maintain stability or to reproduce the present-day climate without the need for the flux corrections used in earlier versions of the model, and some of which are simple scalings of parameters to give the same scientific behaviour as ~~HadCM3~~ HadCM3B at the lower resolution. These differences between ~~HadCM3 and HadCM3L~~HadCM3B and HadCM3BL, which are described below, are generally consistent with work done to optimise the FAMOUS model (Jones, 2003), which has the same ocean resolution as

~~HadCM3L. Note that the Bristol version of HadCM3L is very different from the Met Office version. The Met Office version was mainly used for the early carbon cycle work (Cox et al., 2000) but required significant flux corrections to ensure that the Atlantic surface climate was reasonable. Our version does not require flux correction because of changes in bathymetry described below.~~ HadCM3BL.

4.1.1 North Atlantic Bathymetry: "No Iceland"

As described in Sect. 2.3, care was taken when developing HadCM3 to define the bathymetry of the North Atlantic in order to ensure that the appropriate flow through the Denmark Straits was captured. This flow is lost when the ocean resolution is reduced in ~~HadCM3L~~ HadCM3BL as the channel between Iceland and Greenland becomes less than a single grid-cell wide (on the velocity grid) and thus no flow is permitted. Jones (2003) investigated potential modifications to allow increased heat transport through this region, thus alleviating the unrealistic buildup of sea ice in the Nordic Sea, and concluded that the removal of Iceland was the preferred solution. With this modification, the ~~flux corrections required in previous ocean versions of this resolution (HadCM2) (Johns et al., 1997) to maintain a stable~~ improved meridional overturning circulation ~~are no longer required~~leads to more realistic heat transports in the coupled system and alleviates the need for flux correction.

This change also has a knock on effect on the land surface (and ultimately the atmosphere) in that the ~~2~~ two cells that define Iceland have been removed.

4.1.2 Ocean Vertical Diffusion~~Schemes HadCM3L vs. HadCM3~~

~~Vertical diffusion~~

~~In HadCM3L~~In HadCM3BL, the Richardson Number dependence of the vertical ~~mixing coefficient~~ tracer diffusivity is replaced with a constant background ~~mixing~~ rate, as it is in FAMOUS. Jones (2003) describes problems encountered with FAMOUS in the interaction between the mixed layer and deep vertical ~~mixing~~ diffusion schemes, but was found to have little impact on the solution because of the relatively low resolution.

~~HadCM3L~~ For the calculation of vertical diffusion, HadCM3BL uses a different calculation for the density of seawater from ~~HadCM3. HadCM3~~ HadCM3B. HadCM3B calculates all densities relative to a reference level at the surface using the updated equation of state for seawater of UNESCO (1981). This can result in negative density gradients in the deep ocean and hence a negative Richardson Number, which in turn can produce very high diffusivities at depth which Pacanowski and Philander (1981) was never intended to handle (Rickard, 1999). ~~HadCM3L~~ HadCM3BL instead derives Bryan and Cox (1972) third order polynomials for each 250 m depth span of the ocean (Foreman, 2005) to fit the Knudsen–Ekman equation for the density of seawater and does not produce negative density gradients (Rickard, 1999), but the range of salinities covered may be insufficient for some applications. This choice of diffusion scheme is consistent with that used in FAMOUS.

Ocean diffusion parameters

4.1.3 Ocean Isopycnal diffusion

HadCM3L HadCM3BL uses different coefficients for a number of aspects of the diffusion formulation, as described in Table 1. All of these values are consistent with those used in FAMOUS. In addition, the Visbeck et al. (1997) scheme for the calculation of isopycnal thickness diffusion coefficients, introduced in HadCM3 HadCM3B to improve resolution of currents such as Western boundary currents on the 1.25° grid, is not used in HadCM3LHadCM3BL. Instead, fixed values of the coefficients for surface ocean diffusion, deep ocean diffusion and scale depth are specified, as in FAMOUS.

4.1.4 Solar Penetrative RadiationHadCM3L vs. HadCM3

The ratio In HadCM3 the penetration of solar radiation components (RSOL), which is is represented by a double exponential decay with depth, with coefficients determined from observations. The ratio between the shallower decay and deeper decay exponential is controlled by a parameter called RSOL. This is set to 0.0 in HadCM3 HadCM3B and is set to 3.8×10^{-1} in HadCM3LHadCM3BL, as it is in FAMOUS.

4.1.5 IslandsHadCM3L vs. HadCM3

HadCM3 HadCM3B defines 6 islands in the barotropic solution, around which non-zero depth-integrated flow is permitted: Antarctica, Australia, New Zealand, the Caribbean, Madagascar and Iceland. In HadCM3LHadCM3BL, there is no island for Iceland as this is entirely absent from HadCM3L HadCM3BL and Madagascar is also not defined as an island due to its proximity to Africa.

4.1.6 Sealing and Retuning for Lower Resolution HadCM3L vs. HadCM3

4.2 HadAM3BH

There are a number of further parameters that are scaled or retuned appropriately for the lower resolution ocean to give the same behaviour as HadCM3. These parameters are all scaled in the same manner as for FAMOUS and described in Table 1.

4.3 HadAM3H

HadAM3H HadAM3BH is a higher resolution version of the atmosphere-only variant, HadAM3. HadAM3B. This model is different to that used by the Met Office (e.g., Hudson and Jones, 2002; Arnell et al., 2003) which keeps to 19 levels in the vertical but has some changes to the parameterisations, particularly in the boundary layer. It is used for studies in which the atmospheric circulation is critical, and as such is best represented at high resolution. Its horizontal resolution is three times greater than HadAM3 HadAM3B both latitudinally and longitudinally, i.e., 288×217 288×217 grid points ($1.25°\times0.83°$). The number of vertical levels is increased from 19 to 30, with the extra levels being concentrated close to the Earth's surface and the upper levels remaining similar to HadAM3HadAM3B. The higher spatial resolution requires a smaller timestep of 10

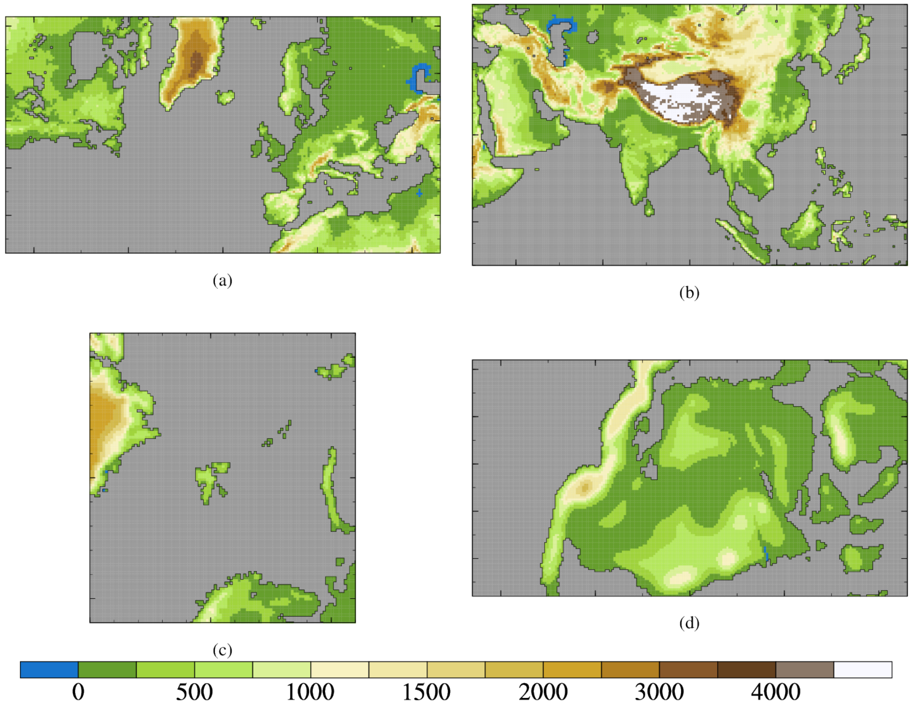


Figure 1. Land/sea mask and orography (sea coloured grey, land height in meters) for four configurations of ~~HadRM3~~HadRM3B. (a) Shows the standard European domain at 0.44° resolution, (b) shows the equivalent domain for E. Asia, (c) shows a configuration for the Arctic and Svalbard at 0.22° (as used in Day et al., 2012), and (d) a N. America/European configuration for the early Cretaceous at 0.44° resolution (as used in Haywood et al., 2004)

minutes. It may be used with either the MOSES1 or MOSES2.1 land surface scheme, and can be used with TRIFFID, though this has rarely been done. The time stepping algorithm is slightly different, in that the dynamics can be updated multiple times between the full physics time steps. In ~~HadAM3H~~HadAM3BH, we use two dynamic per physics time step to allow for improved numerical stability of the model. Various diffusion coefficients, critical relative humidity and parameters for the gravity wave drag scheme have been retuned to account for the change in resolution, as documented in Table 1. Otherwise the model has identical physics to ~~HadAM3~~ HadAM3B and has had no further changes. ~~This model is different to that used by the Met Office (e.g. Hudson and Jones, 2002; Arnell et al., 2003) which keeps to 19 levels in the vertical but has some changes to the parameterisations, particularly in the boundary layer.~~

4.3 ~~HadRM3~~HadRM3B

~~HadRM3~~ HadRM3B is the regional climate model (RCM) version of ~~HadAM3~~ HadAM3B which has been used when representation of high resolution atmospheric processes is important, such as around orography or studying extreme events. It can be configured for any domain size and location has commonly been used for studies over Europe (Jones et al., 1995), the Arctic (Day et al., 2013) and Svalbard (Day et al., 2012), and the East Asian Monsoon region (Bhaskaran et al., 1996). It has also been used to model deep time (Haywood et al., 2004).

The BRIDGE version of HadRM3 is based on the same fundamental physics and model structure of as the Met Office HadAM3, and currently is only available with the MOSES1 land surface scheme. We again do not make any substantial changes to the physical parameterisations so the model is largely identical to HadAM3 except for parameters sensitive to resolution.

Regional climate models require either fixed or time evolving data on the large scale and global atmospheric and ocean response to climate forcings to be provided to them at their lateral (atmospheric) and sea surface boundaries, such as potential temperature and specific humidity. The common experiment set-up, used here, is a one-way nested approach, where no information is fed back into the GCM simulation, but the large scale atmospheric circulation patterns, such as the location of the jet streams, are fed in through the Lateral Boundary Conditions (LBCs). For a RCM to have a "parent" GCM is rare, offering a unique opportunity to investigate the effects of dynamical downscaling without modification (or contradiction) of the physics between the driving GCM and the RCM at the lateral boundaries. LBCs are updated every 6 hours and linearly interpolated for timesteps in between. A 4-grid smoothing is applied to global model data entering the regional model domain. Therefore typically, HadRM3 HadRM3B has been run here using HadAM3 or HadCM3 HadAM3B or HadCM3B to produce the lateral boundary conditions, sea surface temperature, and sea ice concentration data, although there have been experiments using SSTs from HadISST and HadGEM, as well as other models in the CMIP5 experiment to analyse the sensitivity of the model to its boundary conditions.

HadRM3 HadRM3B is run on a standard lat-long grid with the pole rotated so that the centre of the domain of interest lies across the equator within the RCM's grid from of reference (see Fig. 1) to reduce variation in the areas of the grid cells. The timestep of the model is five minutes to maintain numerical stability with the increase in spatial resolution which is commonly $0.44°\times0.44°$ ($\sim$ 50 km$\times$50 km) but has also been run at $0.22°\times0.22°$ ($\sim$ 25 km$\times$25 km). Lateral boundary conditions are typically provided every 3 or 4 6 hours and linearly interpolated to each timestep. The main difference between HadRM3 and HadCM3HadRM3B and HadCM3B/HadAM3 HadAM3B in terms of atmospheric dynamics is in the sub grid scale diffusion applied to the horizontal wind component to prevent the accumulation of energy at the smallest scales and noise (see Table 1). In addition, the parameters which control the proportion of a grid box over which convective and large scale precipitation are assumed to fall, as well as diffusion parameters, vary compared to HadAM3 HadAM3B (see Table 1, variables `conv_eps` and `ls_eps`).

Simulations using the regional climate models have enabled improved spatial representation of temperature and precipitation patterns and response to climate forcings, particularly around mountains and coastlines. The increase in resolution also improves the simulated temporal variability, including simulation of extremes (Durman et al., 2001).

5 Comparison with data

The aim of this section is to qualitatively and quantitatively evaluate the suite of HadCM3@Bristol models in terms of their ability to recreate key aspects of the climate system relative to observations, and other models within the CMIP5 family. In the following subsections, a selection of observational datasets are compared to multiple modelled climatic variables. Details on the datasets used for each variable are briefly outlined in each subsection. This is not intended to be a complete model evaluation; however, it will highlight that some variants do a more realistic job than others at representing various environmental processes. Where appropriate, stronger or weaker models will be highlighted, and some other CMIP5 models will be shown for comparison. Because much of our work at Bristol involves carrying out palaeoclimate or idealised simulations, our standard control simulations are static pre-industrial simulations, similar to the CMIP5 DECK pre-industrial simulation (Eyring et al., 2016a). However, most observational datasets are of the instrumental record, typically the last few decades. This is to be considered when interpreting our evaluation, although it is likely that differences between pre-industrial and the instrumental period are generally small relative to the model biases.

A quantitative evaluation (global RMSE analysis) of the four base state BRIDGE models, namely ~~HadCM3~~ HadCM3B with the MOSES1 land surface scheme, ~~HadCM3~~ HadCM3B with MOSES2.1, ~~HadCM3L~~ HadCM3BL with MOSES2.1 and FAMOUS with MOSES1, is performed against reanalysis and/or observational data and shown alongside new and predecessor models from the CMIP5 database (Fig. 2; BRIDGE models highlighted in red). Here we make use of the ESMValTool(v1.0); a community diagnostic and performance tool (Eyring et al., 2016b) to assess and compare the magnitude of known systematic biases inherent in all climate models. Better understanding of these biases is instrumental in diagnosing their origin and a models ability to reproduce observed spatial and temporal variability and trends in various atmospheric (e.g., Large-scale circulation) and oceanic phenomena (e.g., ENSO). CMIP5 model data is provided from http://www.ceda.ac.uk, while observational (obs4MIPs; Ferraro et al., 2015) and re-analysis (ana4MIPs; Ferraro et al., 2015) data are provided from https://www.earthsystemcog.org, all conforming to the CMIP5 format. Here the BRIDGE models have also been standardised to the CMIP5 format. Further, models and observations are re-gridded to the coarsest resolution within the ESMValTool framework for evaluation. Table 4 details the different metrics used for the evaluation of the historical model simulations in Fig. 2. The BRIDGE models are only pre-industrial climatologies (30-year) without any year-on-year historical forcing, however this is not expected to be detrimental for the evaluation. The results in Fig. 2 demonstrate that the BRIDGE suite of models, with the exception of FAMOUS-M1H, accurately reproduces observed global spatio-temporal patterns. Indeed, ~~HadCM3-M1, HadCM3-M2~~HadCM3B-M1, HadCM3B-M2.1 and in most respects ~~HadCM3L-M2~~HadCM3BL-M2.1, outperform many of the higher fidelity CMIP5 models with lower RMSE when compared to the observations, particularly in respect to global air temperature (at 850 hPa and 200 hPa), U-wind (at 850 hPa and 200 hPa) and 1.5 m surface temperature. It is likely that the course resolution of FAMOUS has a detrimental impact on its performance. The following sections provide a more detailed evaluation of various atmosphere, ocean and land surface variables in the BRIDGE model suite.

Table 4. Observational and reanalysis datasets used for the evaluation in Fig. 2

Performance metric	Obs. dataset	Re-analysis dataset	Year(s) for comparison
TOA outgoing All-sky Short wave radiation (rsut_Glob)	CERES-EBAF	—	2001–2012
TOA outgoing All-sky Long wave radiation (rlut_Glob)	CERES-EBAF	—	2001–2012
Precipitation (pr_Glob)	GPCP-SG	—	1979–2005
Near-surface temperature (tas_Glob)	—	ERA-Interim NCEP	1979–2005 1979–2005
Specific humidity (~~400 hPa~~400 hPa) (hus_Glob-400)	— AIRS	ERA-Interim —	1979–2005 2003–2010
Geopoential height (~~500 hPa~~500 hPa) (zg_Glob-500)	—	ERA-Interim NCEP	1979–2005 1979–2005
V-wind height (~~200 hPa~~200 hPa) (va_Glob-200)	—	ERA-Interim NCEP	1979–2005 1979–2005
V-wind height (~~850 hPa~~850 hPa) (va_Glob-850)	—	ERA-Interim NCEP	1979–2005 1979–2005
U-wind height (~~200 hPa~~200 hPa) (ua_Glob-200)	—	ERA-Interim NCEP	1979–2005 1979–2005
U-wind height (~~850 hPa~~850 hPa) (ua_Glob-850)	—	ERA-Interim NCEP	1979–2005 1979–2005
Temperature (~~200 hPa~~200 hPa) (ta_Glob-200)	—	ERA-Interim NCEP	1979–2005 1979–2005
Temperature (~~850 hPa~~850 hPa) (ta_Glob-850)	—	ERA-Interim NCEP	1979–2005 1979–2005

5.1 Atmosphere

5.1.1 ~~Temperature~~Surface temperature patterns

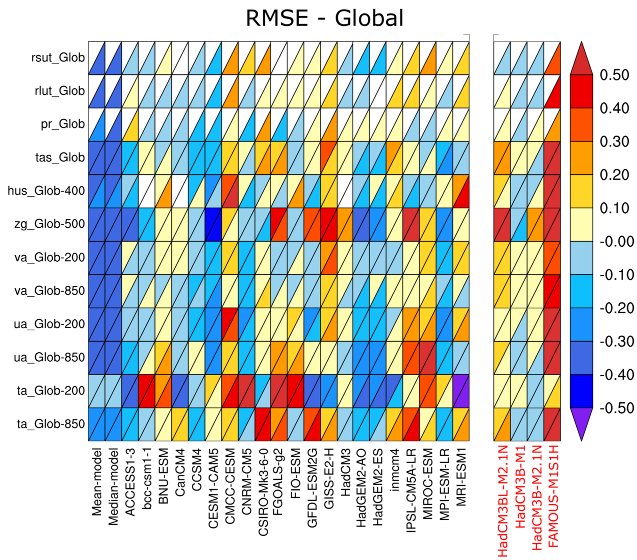


Figure 2. Relative error measure of the CMIP5 models (21 in total; in black) and the BRIDGE models (four in total; in red) performance. Error measure is calculated from a time-space root-mean square error (RMSE) of contemporary and predecessor CMIP5 model historical climatological (1980–2005) seasonal cycle simulations and BRIDGE pre-industrial seasonal-cycle climatologies against observations (1980–2005) for a set of nine different atmospheric variables. Error for each individual variable is characterised as a relative error by normalising the result of the median error of all model results (Gleckler et al., 2008), the BRIDGE models are not included in the mean/median error. For instance, a value of 0.20 indicates that a model~~'~~’s RMSE is 20 % larger than the median CMIP5 error for that variable, whereas a value of −0.20 means the error is 20 % smaller than the median error. The diagonal split grid square shows the relative error for the reference observed/reanalysis dataset (lower right triangle) and the alternative dataset (top left triangle). White triangles/boxes indicate where no data was available. Evaluated global atmospheric variables are TOA outgoing All-sky short wave radiation (rsut_Glob), TOA outgoing ~~Allsky~~ All-sky outgoing ~~Long~~ long wave radiation (rlut_Glob), ~~Precipitation~~ precipitation (pr_Glob), near-surface temperature (tas_Glob), ~~Specific~~ specific humidity at ~~400-hPa~~ 400 hPa (hus_~~Glob- 400~~Glob-400), ~~Geopotential~~ geopotential at ~~500-hPa~~ 500 hPa (zg~~§~~_Glob-500), V-wind at ~~200-hPa~~ 200 hPa (va_Glob-200), V-wind at ~~850-hPa~~ 850 hPa (va_Glob-850), U-wind at ~~200-hPa~~ 200 hPa (ua_Glob-200), U-wind at ~~850-hPa~~ 850 hPa (ua_Glob-850), ~~Temperature~~ temperature at ~~200-hPa~~ 200 hPa (ta_Glob-200), ~~Temperature~~ temperature at ~~850-hPa~~ 850 hPa (ta_~~Glob- 850~~Glob-850).

~~The land surface air temperature (SAT) observation data used here is~~ We compare the modelled temperature and precipitation to observational data provided by the University of East Anglia high resolution climatology for 1960–1990 (CRU CL v2.0) (New et al., 2002). This record is based on a range of weather stations totalling more than 10 thousand stations for temperature and more than 25 thousand stations for precipitation, with the best spatial coverage over North America, Europe and India and the sparsest spatial coverage over the interiors of South America and Africa and Antarctica. Modelled SAT fields were masked to model land points only and differences to observations were done at the same resolution as the relevant model, as shown in Fig. 3.

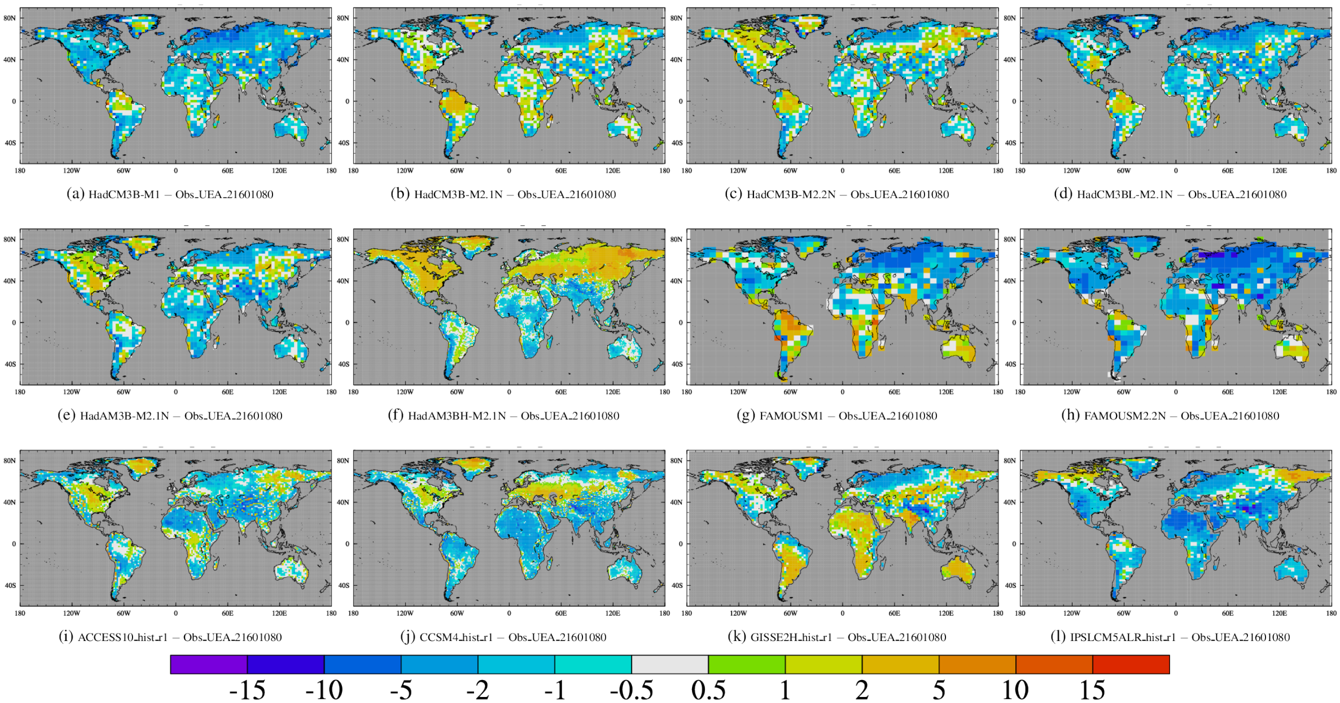


Figure 3. (a) The difference between the annual mean surface air temperature (in °C) of ~~our version of HadCM3-M1~~ HadCM3B-M1 and the CRU CL v2.0 for the period 1960–1990 regridded onto the ~~HadCM3-M1~~ HadCM3B-M1 grid, (b) As (a) for the ~~HadCM3-M2~~HadCM3B-M2.1 version, (c) As (a) but for ~~HadCM3-M2~~HadCM3B-M2.2, (d) As (a) but for the ~~HadCM3L-M2~~HadCM3BL-M2.1N version, (e) As (a) but for ~~HadAM3-M2~~HadAM3B-M2.1, (f) As (a) but for ~~HadAM3H-M2~~HadAM3BH-M2.1, (g) As (a) but for FAMOUS-M1, and (h) As (a) but for FAMOUS-M2.2. (i), (j), (k) and (l) show comparable results for four CMIP5 models, ACCESS1-0, CCSM4, GISS-E2-H and IPSL-CM5A-LR respectively. These were chosen to represent two models which were above the CMIP5 average in terms of their RMSE with respect to surface air temperature, and two models which were below average. All differences are calculated by regridding the CRU data onto the corresponding model grid, using simple bi-linear interpolation.

It should be noted that the comparison between the versions of the ~~HadCM3~~ HadCM3B family and the observed CRU CL v2.0 data is not a "clean" comparison. The observed data is for 1960–1990 whereas all model simulations are for the pre-industrial period. In the case of ~~HadAM3~~ HadAMB3 simulations, the SSTs used are the 1870–1900 means of HadISST. To evaluate the impact of this effect, we examined the CMIP5 historical ~~run of~~ experiment of HadCM3-M1 done at the Hadley Centre ~~version of HadCM3-M1~~((Smith et al., 2007, 2010)). The differences between the 1960–1990 climate means compared to the 1860–1890 climate means were generally small compared to the model biases, with the overall mean warming between the two periods being 0.6 °C. Similarly, the four CMIP5 simulations are averages from 1860–1890 of the historical runs (using one ensemble member only, r1i1p1) and so the comparisons to the ~~HadCM3~~ HadCM3B family are not perfectly clean.

~~HadCM3-M1~~ HadCM3B-M1 (Fig. 3a) generally has a small cold bias compared to the data, with most regions experiencing colder temperatures ~~of~~ by 2 °C to 3 °C. The area weighted root mean square differences (RMSE) is 2.8 °C, but with smaller errors in the tropics and a small warm bias in South America. There is also a small warm bias over Greenland but this should be treated with some caution since there are issues about elevation effects and the data is relatively sparse in this region. The results for Fig. 3a, are largely identical to those calculated using the CMIP5 HadCM3-M1 archived data (run by the UK Hadley Centre) for the historical run averaged between 1860–1899 inclusive (not shown). The differences are mostly less than 0.5 °C and never exceed 1 °C, with an RMSE of 0.5 °C. Differences between the 1860–1889 average and the 1960–1989 average for the CMIP5 historical run are small, verifying that the model biases greatly exceed any differences between pre-industrial and modern temperatures. However, the small warming that does occur between 1860–1889 and 1960–1989 does reduce the cold bias marginally (RMSE decreased by 0.1 °C).

Using MOSES 2, ~~HadCM3-M2~~HadCM3B-M2.1 (Fig. 3b) shows a significant reduction in the cold bias, resulting in a RMSE of 2.1 °C. The cold bias has reduced but still remains over northern Russia and Scandinavia, while over South America (Amazon) and Greenland the warm anomalies have intensified. Over the Amazon this is likely due to the difficulties in the vegetation model (see Sect. 5.3.1), while difficulties with Greenland were mentioned above. Elsewhere, the general cool bias seen in Fig. 3a has gone, replaced by anomalies of ±2 °C to 5 °C, with few widespread regional anomalies. Similarly, ~~HadCM3-M2~~HadCM3B-M2.2N (Fig. 3c) also shows a reduced cold bias, with an RMSE of 2.1 °C. This model variant shows a slight reduction in the warm anomaly observed over the Amazon compared to Fig. 3b, but has a more extensive warm bias of 1 °C to 2 °C at higher northern latitudes, e.g., over North America.

~~HadCM3L-M2~~HadCM3BL-M2.1 (Fig. 3d) has a RMSE of 2.6 °C and a comparable cold bias to ~~HadCM3-M1~~HadCM3B-M1. As with the ~~HadCM3~~ HadCM3B model variants, using MOSES2 with ~~HadCM3L~~ HadCM3BL reduces the cold bias and RMSE compared to using MOSES1, with ~~HadCM3L-M1~~ HadCM3BL-M1 having a much higher RMSE (not shown). Once again, the high northern latitudes (particularly over Russia and Scandinavia) are too cold, which is the result of an exaggerated seasonal cycle due to an overly cold winter. This is also the case for other ~~HadCM3~~ HadCM3B model variants, but it is most pronounced for the ~~HadCM3L~~ HadCM3BL variants. Similarly to the other simulations using MOSES 2, the Amazon remains slightly warmer than the observations with slightly reduced broadleaf forest cover (see Sect. 5.3.1).

The atmosphere-only models vary significantly depending on their resolution. At standard resolution, ~~HadAM3-M2~~HadAM3B-M2.1 (Fig. 3e) shows similar spatial anomalies and RMSE to Figs. 3a–d, but greater warm biases over North America and Greenland of up to 5 °C and cool biases over Africa and southern Asia of 2 °C to 5 °C. However, it has the smallest anomaly over the Amazon compared to the other standard resolution model variants, and a comparable RMSE (2.3 °C). ~~HadAM3H-M2~~HadAM3BH-M2.1 (Fig. 3f) on the other hand shows a markedly different spatial pattern in its temperature biases to the model versions already described. It is the only simulation not to show a global cold bias. This is due to warmer than observed temperatures of 2 °C to 5 °C over the majority of land surfaces north of 30° N (with the exception of the southern tip of Greenland and mountainous regions). It has a slight cold bias of 1 °C to 2 °C over areas south of 30° N (with the exception of some regions in South America). Although these biases are extensive spatially, they are not of greater magnitude than the regional biases found in other model variants or CMIP5 models and the RMSE of the ~~HadAM3H-M2~~HadAM3BH-M2.1 simulation is 2.2 °C.

The FAMOUS model variants (Figs. 3g and 3h) have larger RMSE values than the higher resolution model variants and the other CMIP5 models. FAMOUS-M2.2 (Fig. 3h) is the worse of the two, with a RMSE of 4.1 °C and extreme cold biases over Northern Hemisphere continents, which exceed 10 °C around Scandinavia. The cold bias in Fig. 3g is less extreme, but instead has a warm bias in South America of up to 5 °C to 10 °C and up to 2 °C to 5 °C over India and Australia. There is some improvement in the RMSE for FAMOUS-M1, but it is still much higher (3.3 °C) than the higher resolution model variants.

Some of the differences between the mean annual temperature biases in the models are due to changes in the models themselves. For instance, the improvement generally seen between models with MOSES1 and MOSES2.1 are primarily due to the better representation of the land surface, particularly the snow cover, as discussed above. It's also notable that the lower resolution ocean models tend to be cooler in the higher latitudes, as the lower resolution ocean makes it more difficult to move heat away from the equator.

For comparison, we show the SAT fields from four CMIP5 models (Figs. 3i–l), selected based on the results of the IPCC AR5 WG1 model evaluation (Chapter 9). We selected two models which were above average for their simulation of SAT (ACCESS1-0 and CCSM4) and two models which were below average (GISS-E2-H and IPSL-CM5A-LR). In all cases these models are not the best or worst extremes, but represent the typical range of model skill. Again, the observations have been interpolated onto the appropriate resolution of the model from which the RMSE was calculated. As can be seen, the general picture that emerges is that most of the varieties of ~~HadCM3~~ HadCM3B (except perhaps for FAMOUS) are well within the skill of the CMIP5 ensemble. The CMIP5 models all show large regional biases of up to ±5 °C (with little consistency on the sign of the anomaly between them) and the RMSE scores range from 2.3 °C to 3.3 °C, which are similar to the varieties of the ~~HadCM3~~ HadCM3B model. Indeed ~~HadCM3-M2~~HadCM3B-M2.1N and ~~HadCM3-M2~~HadCM3B-M2.2 have the smallest RMSE values of the models sampled.

5.1.2 Precipitation

Figure 2 shows that the BRIDGE models with the exception of FAMOUS produce annual precipitation amounts comparable to other CMIP5 models suggesting that our models are capturing the general synoptic scale features (frontal, convective and mesoscale).

While global annual RMSE for the BRIDGE models compare favourably it is also key to investigate the mean spatial patterns of precipitation to ascertain whether the models are reproducing these patterns in accordance with the observations. We assess annual climatological precipitation for the BRIDGE model suite against CRU CL v2.0 (New et al., 2002), a high resolution (0.5°×0.5°) global land surface product (excluding Antarctica). The resolution is transformed (bi-linear interpolation) to the appropriate grid in the model. We are again comparing our pre-industrial simulations with 1960–1990 observations, but the model biases are generally much larger than any trends.

Figure 4 shows the regional biases in mean annual precipitation, expressed as a % error compared to the CRU CL data. For consistency with the previous figure, we also include the same four CMIP5 models. Regionally, spatial patterns in precipitation bias are generally consistent between the different BRIDGE models and broadly comparable to their CMIP5 models.

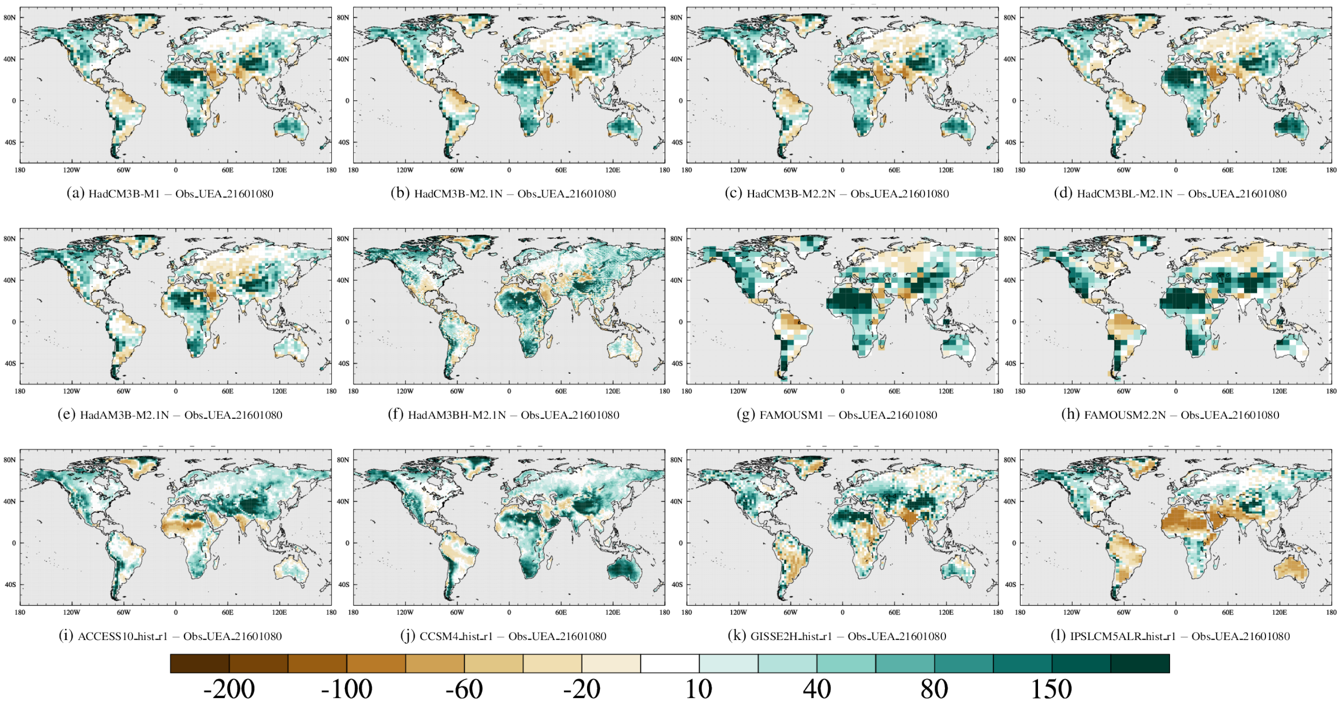
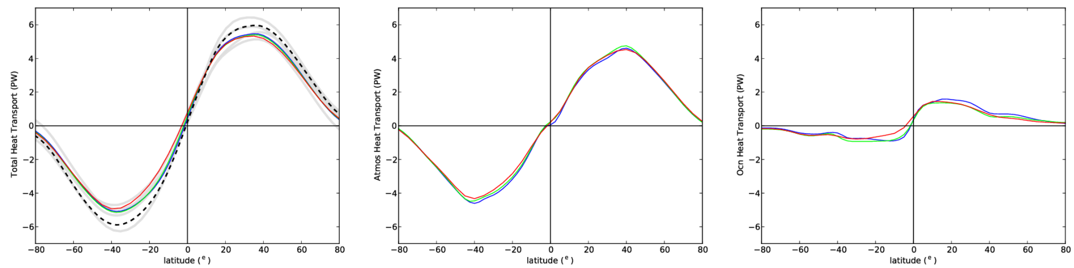


Figure 4. As for Fig.3 except showing the difference in mean annual precipitation, expressed as a % difference to the CRU CL v2.0 observations.

The BRIDGE simulations ~~affirm a known problem in all~~ have a similar problem to many CMIP5 models whereby they overestimate precipitation in regions of ~~steep~~ topography. This is particularly noticable around the Himalayas and Tibet but is also visible on the upstream side of the Rockies and Andes. This may be due to poor representation of moisture gradients and regional dynamics~~, but~~. However, the apparent discrepancy with observations can be amplified by a known negative bias in gauge stations ~~amplifying model-observation disparity (Adam et al., 2006). Secondly, they~~ (Adam et al., 2006). Gauge stations also underestimate precipitation leeward of ~~the mountain range~~ mountain ranges (e.g., ~~Himalaya's~~ Himalayas and Andes), as well as over arid regions, ~~again a known bias in the CMIP5 model suite~~which can contribute to model-data discrepancy in these regions also.

Monsoonal regions of south-east Asia, Austral-Asia, southern South America, West and central Africa, overestimate precipitation by 0.5 mm day^{-1} to 2 mm day^{-1} while underestimate precipitation in the Indian and northern arm of the South American region by $\sim$ 1 mm day^{-1} to 4 mm day^{-1}. There are however some exceptions with ~~HadAM3-M2~~HadAM3B-M2.1 (Fig. 4), an atmosphere-only GCM using observed SST (HadISST 1870–1900) producing a more reasonable precipitation signal compared to the observations suggesting the importance in the accuracy of SST/local ocean circulation dependency (this is also seen in Australia). There is still a problem with the ITCZ location over South-America being too far South giving this north-south dipole in negative/positive anomalies.



(a) Total (b) Atmosphere (c) Ocean

Figure 5. The annual mean northward heat transport in total, the atmosphere and the ocean. Black line shows the observational estimate, blue: ~~HadCM3-M2~~HadCM3B-M2.1, green: ~~HadCM3L-M2~~HadCM3BL-M2.1 and red: FAMOUS-M2.2, grey lines show transports for a selection of CMIP5 models. These transports are calculated as the implied heat transports from the TOA and surface energy fluxes (see Trenberth and Fasullo (2013) for details). Observational estimates for the total transport are derived from the CERES data.

It is also noted that an increase in resolution does not produce a noticeable improvement on spatial annual precipitation bias in certain monsoon regions in the BRIDGE models (Fig. 4f compared to 4j) again suggesting the importance of accuracy in SST and ocean circulation, with the exception of South America where there is improvement. Spatially, increased resolution does not ~~effect~~ affect the sign of anomaly nor the spatial patterns of precipitation regionally (with the exception of South America) throughout the BRIDGE suite of models however the magnitude of the precipitation bias does progressively decrease.

5.1.3 ~~Top of the atmosphere radiation fluxes~~Horizontal heat transports

There is broad agreement between the observed and simulated total northward heat transport. Similarly, the partitioning between the ocean and atmosphere is qualitatively similar to that estimated by Trenberth and Fasullo (2013). We find that all versions of the model simulate heat transport that are consistent with CMIP5 models (see grey lines in Fig. 5). However, in common with almost all other climate models we find that on the equator, although the total heat transport is northward, in agreement with the observations, the atmospheric heat transport is also northward, contrary to the observed southward transport (Loeb et al., 2016). The cause of this in any of the models in which it is a feature is unclear. The three versions of ~~HadCM3~~ HadCM3B show remarkably similar amounts of total heat transport, the major difference is FAMOUS which underestimates the southward heat transport in the Southern Hemisphere subtropics rather more than ~~HadCM3 and HadCM3L~~HadCM3B and HadCM3BL. This is due to the smaller amount of ocean heat transport in this region in FAMOUS. This discrepancy is not due to the coarse resolution of the FAMOUS ocean because, interestingly, in this region the ocean heat transport in ~~HadCM3L~~ HadCM3BL is very similar to ~~HadCM3~~ HadCM3B whose ocean resolution is quite different. Therefore it is more likely that the difference arises from the atmospheric forcing of the surface ocean. In the Northern Hemisphere the ~~HadCM3L~~ HadCM3BL ocean heat transport is more similar to ~~AMOUS~~FAMOUS, suggesting that the ocean resolution is more important here. This is likely due to the processes that determine the ocean's overturning circulation being simulated rather differently in the higher and lower resolution models.

5.2 Ocean

5.2.1 Sea Surface Temperature

The BRIDGE suite of models is capable of reproducing the broad global latitudinal patterns and gradients in SST (Fig. 6). Nonetheless, some cold and warm biases of over 8 °C are present especially where sharp fronts and boundary currents are not resolved. Other biases of similar magnitude also appear in the upwelling regions (e.g., west of Africa and of South America), and again these are likely associated with processes that are not fully resolved by the model. Colder SSTs in the sub-polar North Atlantic for all models are not uncommon and likely due to the coarse resolution (e.g., Marzocchi et al., 2015a). This can be seen by comparing ~~HadCM3 and HadCM3L~~HadCM3B and HadCM3BL, which are models that differ most in their ocean resolution. Cold biases in the Northern Hemisphere are more extensive in ~~HadCM3L than in HadCM3~~HadCM3BL than in HadCM3B. Warmer SSTs of up to 8 °C are present in the Southern Hemisphere, especially in the Southern Ocean, both in ~~HadCM3 and HadCM3L~~HadCM3B and HadCM3BL. FAMOUS is characterised by colder than observed SSTs in the Northern Hemisphere, in common with ~~HadCM3 and HadCM3L~~HadCM3B and HadCM3BL, and warmer SSTs by up to 8 °C almost everywhere in the Southern Hemisphere ~~. HadCM3 and HadCM3L~~ despite the bias in ocean heat transport. HadCM3B and HadCM3BL do not show any notably larger biases when compared to typical CMIP5 models. All of the ~~HadCM3~~ HadCM3B models, including FAMOUS, show smaller temperature biases in the Southern Ocean than GISS-E2-H, and the biases in the North Pacific are of a similar magnitude to those in IPSL-CM5A-LR.

5.2.2 Sea Surface Salinity

The broad global latitudinal patterns of sea surface salinity are realistically reproduced by the suite of BRIDGE simulations (Fig. 7). However in the global average, the models show a fresh bias of about 0.5 $\mathrm{g\,kg^{-1}}$, as we shall show in the following section this is likely related to the rather different vertical structure of the ocean in the model than in the observations. In all models, substantial differences from the observations are found in the Arctic Ocean, exhibiting higher salinities (up to 10 $\mathrm{g\,kg^{-1}}$) in the Kara Sea and generally north of Russia. Generally lower salinities (of up to 5 $\mathrm{g\,kg^{-1}}$) are found in the Chukchi and Beaufort seas. The largest differences are found in enclosed or semi-enclosed basins, such as the Mediterranean Sea, where it is more saline, or the Black Sea, Caspian Sea and Hudson Bay where it is markedly fresher. In all versions of the model the subtropical North Atlantic tends to be more saline than the observations.

Substantial differences from the observations can also be found in CMIP5 models (Figs. 7g–h) with magnitudes comparable to the BRIDGE models. We note that some of the differences at high latitudes could be due to biases in the simulation of sea ice concentration and distribution.

5.2.3 The Atlantic Meridional Overturning Circulation

Figure 8 shows the mean strength of the Atlantic Meridional Overturning Circulation (AMOC) for the three main model families (~~HadCM3, HadCM3L~~ HadCM3B, HadCM3BL and FAMOUS). Values are shown as zonally integrated depth profiles

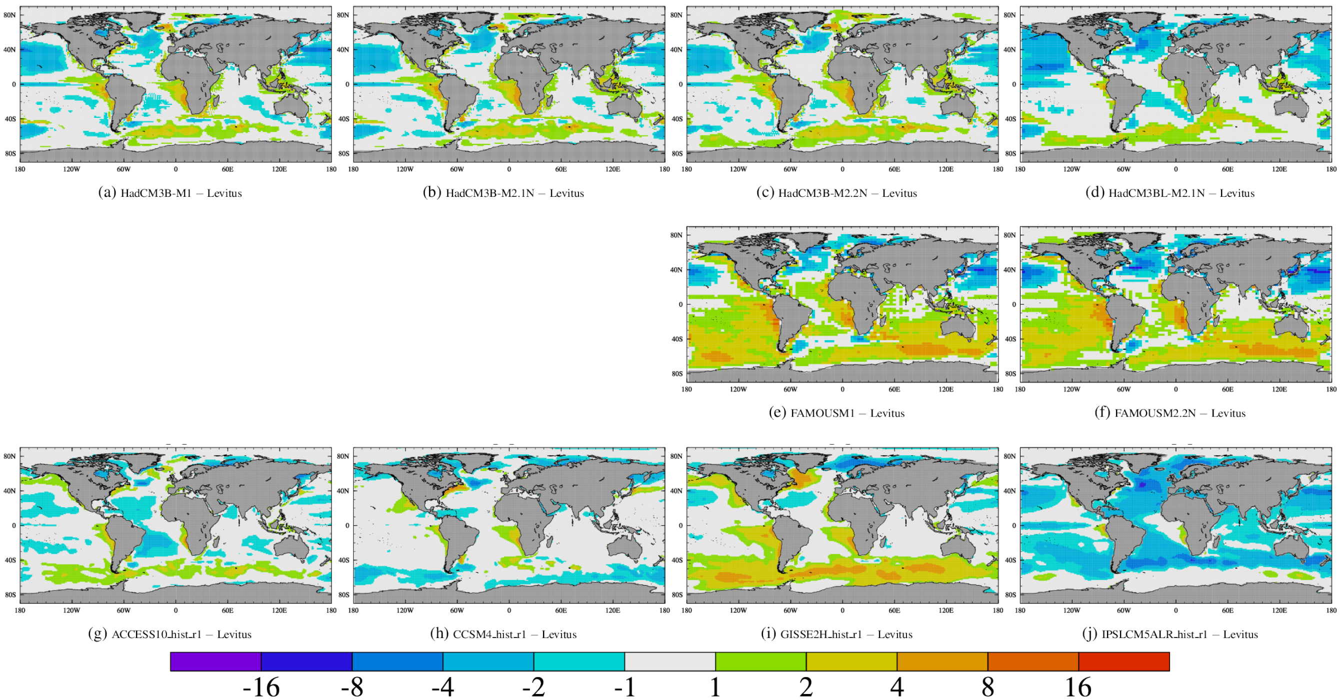


Figure 6. Annual mean sea surface temperature differences (in °C) for a range of coupled model simulations, and also for the same four CMIP5 models used in Figs. 3 and 4. The observational dataset is the Levitus World Ocean Atlas (2009) (Locarnini et al., 2010). The figure shows the difference in SST between model and observations for (a) ~~HadCM3-M1~~HadCM3B-M1, (b) ~~HadCM3-M2~~HadCM3B-M2.1, (c) ~~HadCM3-M2~~HadCM3B-M2.2, (d) ~~HadCM3L-M2~~HadCM3BL-M2.1, (e) FAMOUS-M1, (f) FAMOUS-M2.2, (g) ACCESS1.0, (h) CCSM4, (i) GISS-E2-H, and (j) IPSL-CM5A-LR. Model output is regridded to the same resolution of the observations.

measured in terms of the northward flow of water at 26.25° N. The modelled AMOC is compared to observations from the Rapid Climate Change-Meridional Overturning Circulation and Heatflux Array (RAPID-MOCHA) at 26.5° N (Smeed et al., 2015), which have been calculated from daily data spanning the 2nd April 2004 to the 30th March 2015.

The strength of the AMOC varies on an annual basis so a range of values is shown for both the models and observations, with the depth at which the AMOC peaks highlighted with a point. The peak flow of the North Atlantic Deep Water (NADW) cell identified by the RAPID-MOCHA array lies at around 1000 m and varies ~~annually~~ from year to year between 14 Sv and 19 Sv. All three models do a reasonable job at modelling the NADW cell in terms of the magnitude of maximum flow. However, maximum overturning is too shallow for all model variants, peaking at approximately 800 m. ~~HadCM3L has a larger annual variation,~~ HadCM3BL shows larger year to year variability than the observations: approximately twice as large as that ~~of the observations, resulting in an underestimation of the annual minimum peak northward transfer~~in the observations. This results in years with a lower minimum volume transport than are seen in the observations. FAMOUS model variants tend to underestimate the ~~annual~~ year to year variation by approximately 50 %, ~~while HadCM3~~ although this is in contrast to the study

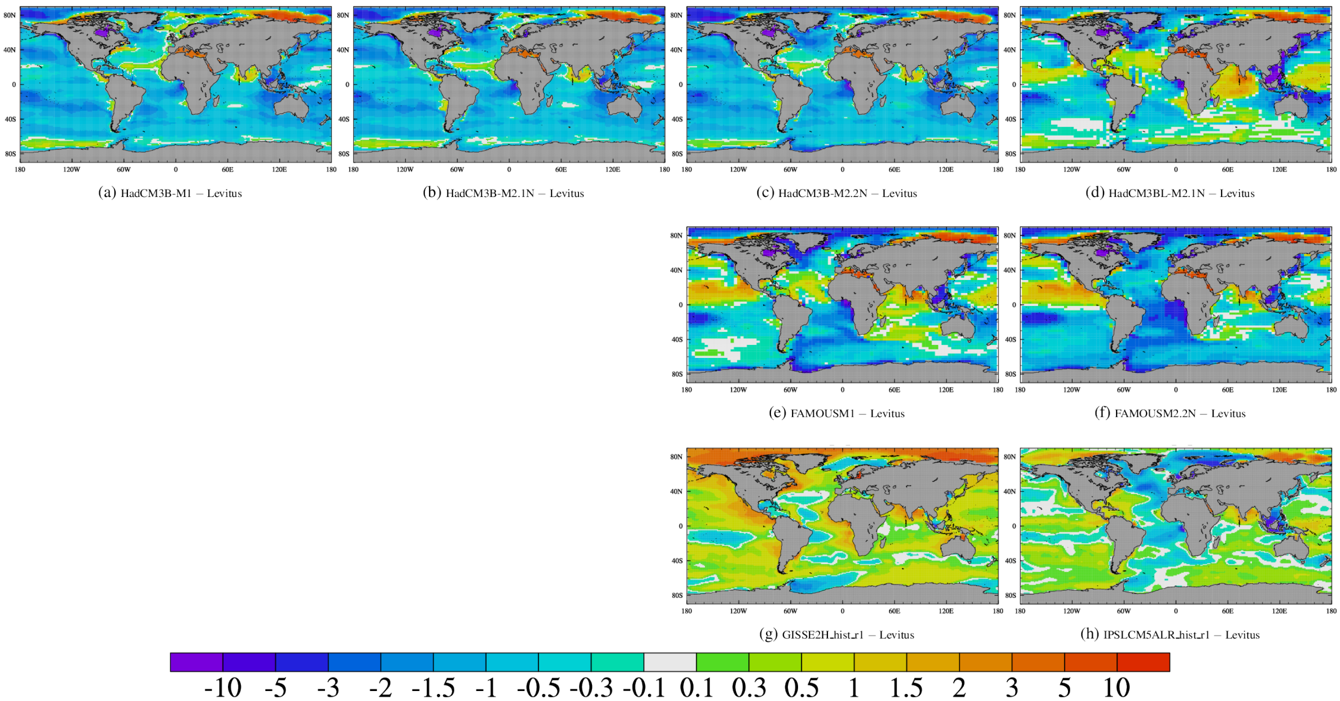


Figure 7. As Fig. 6 but showing the differences in sea surface salinity (in $\mathrm{g\,kg^{-1}}$) between models and observations. (a) ~~HadCM3-M1~~HadCM3B-M1, (b) ~~HadCM3-M2~~HadCM3B-M2.1, (c) ~~HadCM3-M2~~HadCM3B-M2.2, (d) ~~HadCM3L-M2~~HadCM3BL-M2.1, (e) FAMOUS-M1, (f) FAMOUS-M2.2, (g) GISS-E2-H, and (h) IPSL-CM5A-LR. The observational dataset is the Levitus World Ocean Atlas (2009) (Antonov et al., 2010). Model output is regridded to the same resolution of the observations.

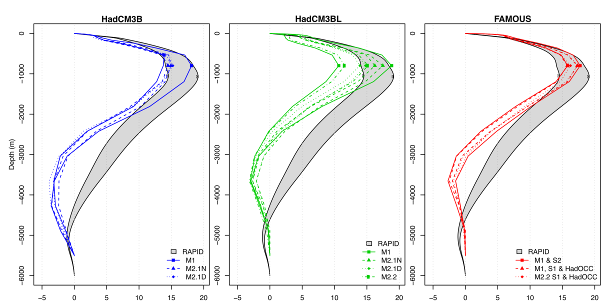


Figure 8. Annual depth profiles of the Atlantic meridional overturning circulation (AMOC) at 26.25° N showing range of values for variants of ~~HadCM3~~HadCM3B, ~~HadCM3L~~ HadCM3BL and FAMOUS. Annual data from the RAPID array at 26.5° N is highlighted in grey. The depth at which the AMOC reaches its maximum is indicated with a point.

of Sarojini et al. (2011) who showed that FAMOUS exhibited greater short-term variability than the RAPID-MOCHA array. HadCM3B variants have a realistic ~~annual variation~~ year to year variability at least in the upper 1500 m of the ocean.

All of the models do a poor job at representing the flow of the NADW cell below 2000 m depth. McCarthy et al. (2012) showed that at this latitude, approximately 60 % of the southward return flow is comprised of upper NADW (between 1100 m and 3000 m) and 40 % of the lower NADW (between 3000 m and 5000 m). The modelled stream functions show that the return flow is shifted to shallower depths, indicating a shallower overturning in all of the model variants.

The CMIP5 models exhibit a wide spread in the mean strength of the AMOC, ranging from 13 Sv to 31 Sv and peaking at latitudes between 20° N and 60° N (e.g., Zhang and Wang, 2013). It was not possible to include the CMIP5 models in Fig. 8, however, the studies of Roberts et al. (2013) and Msadek et al. (2013) produced similar plots of AMOC zonally integrated depth profiles for a range of models compared to observations (their Figures 1 and 3 respectively). The ~~HadCM3~~ HadCM3B and FAMOUS variants are shown to have very similar stream function profiles to GFDL Climate Model 2.1, NCAR CCSM4 models and the MPI models, and more accurately simulate the maximum overturning than the NorESM1 model variants. A similar pattern of biases is apparent in the vertical structure for these models, i.e., a too shallow overturning cell, however the point of maximum overturning is shallower in the ~~HadCM3~~ HadCM3B and FAMOUS variants.

This bias in the vertical structure has been attributed in some studies to inaccurate transport in the Nordic Sea overflows, which in the case of ~~HadCM3~~ HadCM3B includes a greater than observed overflow across the Denmark Strait, in addition to subgrid scale processes (see Legg et al., 2009; Roberts et al., 2013).

An additional cause of the shallow overturning may be the excessive surface salinity in the North Atlantic in all model versions, particularly around the subtropics as shown in Fig. 7. The study of Pardaens et al. (2003) investigated the freshwater budget in ~~HadCM3~~HadCM3B, concluding that in the North Atlantic saline conditions are primarily a result of excessive evaporation. Other components, such as insufficient subtropical runoff from the west coast, may also have an influence. This results in the Atlantic being too stratified and consequently too stable which may reduce the depth of overturning.

A further consequence of this is a net northward transport of fresh water into the Atlantic (Liu et al., 2014), which ~~is thought to~~ may result in a monostable stability regime of the AMOC in ~~HadCM3~~ HadCM3B instead of a bistable regime (Weaver et al., 2012; Liu et al., 2014). In contrast, Hawkins et al. (2011) have demonstrated a bistable regime in FAMOUS. Approximately 60 % of the CMIP5 models have been shown to exhibit monostability (Weaver et al., 2012). However this is contrary to what is indicated in ~~observations~~the palaeorecord and inferred from the measurements of diagnostic indicators in the present-day ocean; that there is a net export of freshwater from the Atlantic and consequently the AMOC ~~is~~ may be in a bistable regime. This indicates that the AMOC may be artificially stable in the ~~HadCM3~~ HadCM3B and FAMOUS model variants in addition to a range of other CMIP5 models. There remains uncertainty over this hypothesis however, with Sijp (2012) concluding that freshwater export may not be a reliable indicator of AMOC stability.

5.3 Land

5.3.1 Vegetation Distribution

These models have a simple representation of terrestrial vegetation, with five plant functional types that each covers a large climatic range. Comparing the dominant PFT in the model to a reconstruction of pre-industrial vegetation (Ramankutty and

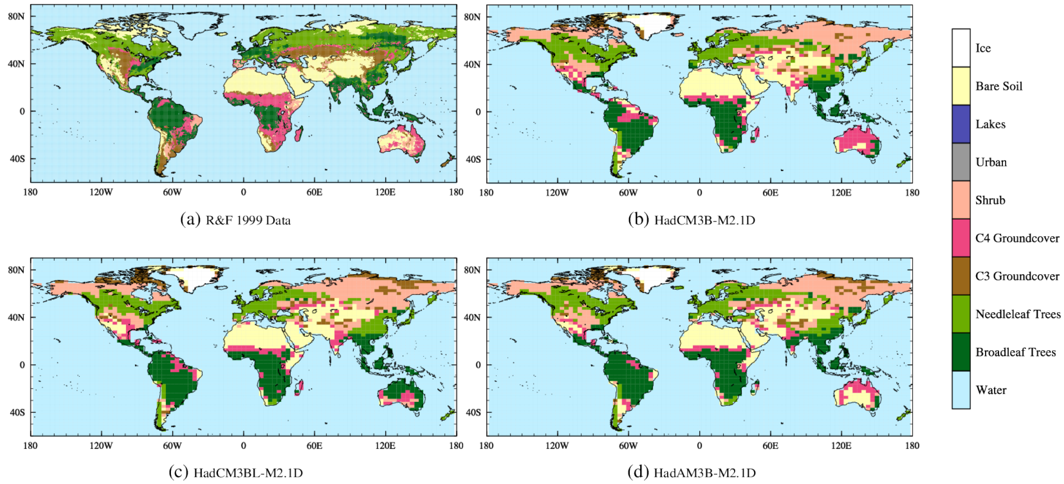


Figure 9. Maps of the dominant plant functional type for observations (a) and model simulations of the pre-industrial. The models shown are (b) HadCM3B-M2.1D (c) HadCM3BL-M2.1D and (d) HadAM3B-M2.1D. The observed dataset for comparison is Ramankutty and Foley (1999).

Foley, 1999), we can see the model captures the overall correct pattern (Fig. 9), with slight errors of extent and/or exact location. Previous studies (Betts et al., 2004) which compared TRIFFID PFT distributions to the IGBP-DIS land cover dataset (which represents the modern distribution of vegetation as derived from satellite image interpretation, Loveland and Belward (1997) found much of the same patterns.

Maps of the dominant plant functional type for observations (a) and model simulations of the pre-industrial. The models shown are (b) HadCM3-M2.1D (c) HadCM3L-M2.1D and (d) HadAM3-M2.1D. The observed dataset for comparison is Ramankutty and Foley (1999).

The broadleaf trees in the models have a known problem of extending tropics tend to extend too far, especially in the Southern Hemisphere, and this as can be seen in FigFigs. 9b–d. The southern mid-latitudes are difficult to capture accurately, for a variety of reasons, including the challenge of precipitation patterns in this region. The HadCM3L HadCM3BL model is significantly worse than either HadCM3 or HadAM3 HadCM3B or HadAM3B in this regard, . This is because of its decreased ocean resolution, which affects the sea surface temperature and therefore the water transport to the Amazon region.

A feature which is isolated to this version of appears in the HadCM3 and HadCM3 models is a tendency for the Amazon broadleaf forest extent to be underestimated at the mouth of the River Amazon, even at relatively low carbon dioxide concentrations . (Figs. 9b–d compared to Fig. 9a). At higher carbon dioxide levels, this is a known feature of the model (Malhi et al., 2009; Betts et al., 2004) caused by ocean circulation resulting in insufficient precipitation to sustain the forest. Since the model underestimates precipitation , this The tendency of the coupled models to underestimate precipitation in this area is apparent in

Fig. 4 and is particularly notable in HadCM3. This leads to TRIFFID modelling the presence of C4 grasses instead of broadleaf trees.

Grasses tend to be globally slightly underestimated with the position of vegetation in the Sahara desert and other arid regions well reproduced, but the density is modelled to be too sparse, particularly in south-west Africa, central and south-west Asia, south-west North America and Australia.

The shrub PFT is overestimated at high latitudes, perhaps as a result of the high latitude cold bias in the model. We can see this in these simulations (Figs. 9b–d compared to observations Fig. 9b). There is the correct pattern of needleleaf trees , but lower amounts a). The models simulate less needleleaf trees than observations for Ramankutty and Foley (1999). However, in Fig. 9e, we can see that the models consistently overestimate , instead simulating shrubs. The models also understimate the amount of high northern latitude shrubs, mainly because of underestimation of bare ground (Fig. 9f). grasses and bare soil.

The observational dataset is re-gridded from the original to the nine surface types in our models, which introduces more uncertainty. In particular, the dominant PFT obviously is a difficult metric to consider precisely, as it does not represent mixed vegetation systems such as Savannah, well. Some difficulties mainly originate in how areas such as tundra are allocated – to bare soil or to C3 grasses. Because of the limited number of PFTs in the model, C3 grass represents a large range of low-lying vegetation types, arguably also encompassing mosses and lichen and very sparse tundra vegetation.

There are also some uncertainties associated with the Ramankutty and Foley (1999) dataset, which is a reconstruction of pre-industrial vegetation. Other model-observations discrepancies have been suggested to be a combination of orographic representation leading to underestimation of precipitation and the inadequate treatment of natural disturbance mechanisms such as fire (Betts et al., 2004).

Though not shown, the Equilibrium (run for 50 years every five years) simulations are very similar to the Dynamic (run ever ten days) ones, especially in the tree PFTs. That the Equilibrium and Dynamic simulations from the same model are very closely related suggests that although the inter-annual variability does have some influence on the vegetation, in general the mean climate is more important.

5.3.2 Net Primary Productivity

The Net Primary Productivity (NPP) of the models, compared to MODIS 2001 NPP observations, is good at capturing the global latitudinal patterns, with higher NPP in the tropics and lower in other regions (see Fig. 10). One notable exception is the failure of the model to capture sub-tropical spikes in productivityand the peak for the Amazon tropics , especially at around 20° N, which is also underestimated in the CMIP5 models analysed here (shown by grey lines). The HadCM3B productivity peak over the Amazon tropical area is lower in the model than observations. HoweverOverall, the NPP performance of our models compares favourably with that of CMIP5 models. The large range of NPP values of these CMIP5 models encompasses our models at nearly all latitudes.

The Amazon forest extends a little too far south in all the models, but this is a key area of difference as well, with HadAM3 HadAM3B models better capturing the observed distribution, and the lower resolution ocean of HadCM3L HadCM3BL suffering the most from excess tropical forest. However, HadCM3L HadCM3BL models do better in the Southern Hemisphere,

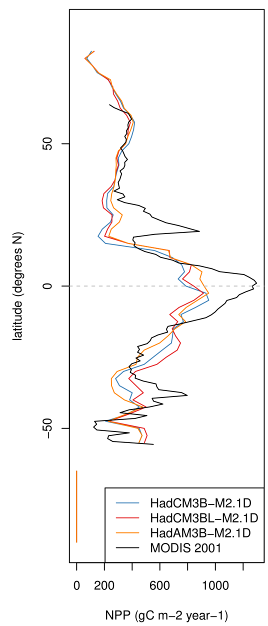


Figure 10. The latitudinal average NPP in $g_C\ m^{-2}\ year^{-1}$. CMIP5 models without dynamic vegetation plotted here are: CCSM4 and IPSL-CM5A-LR. CMIP5 models with dynamic vegetation plotted here are: MIROC-ESM and MPI-ESM.

and better than ~~HadCM3~~ HadCM3B in other parts of the tropics. As in the case of the PFT distribution (upon which the NPP is based) there is a close relationship between the equilibrium (not shown) and dynamic simulations of NPP.

6 Summary and Future Directions

This paper provides an overview of a variety of versions of the HadCM3 family of coupled climate models used in BRIDGE at the University of Bristol. In this study we have termed the BRIDGE variants HadCM3B, in order to distinguish our versions from those originally developed at the Met Office. We provide updated documentation of these variants, including atmosphere-only, low-resolution ocean, and high-resolution atmosphere-only models, and including three alternative versions of the MOSES land surface scheme. Using an up-to-date set of observational benchmarks we show through detailed comparisons, that the models provide a good representation of large-scale features of the climate system, both over land and for the ocean~~and which remains~~. We additionally show that they remain comparable to most CMIP5 models.

The speed and relative complexity of ~~HadCM3~~ HadCM3B and its variants creates opportunities for tackling a range of problems. Large ensembles are possible because of the relatively small number of processors required. Ensembles can explore probabilistic approaches to climate change quantification, model parametric uncertainty or boundary condition uncertainty. Long integrations of many millennia are also possible, so that longer term climate changes, for example covering the last deglaciation, can be investigated.

Several versions of the model are under continued development and improvement. For example, FAMOUS has been coupled to an interactive ice-sheet model (Gregory et al., 2012) to allow predictions of sea-level and land ice on longer timescales. Further developments in this approach will allow more detailed investigation of ~~the~~ climate – sea-level interactions for a variety of times in the past (e.g., Roberts et al., 2014). FAMOUS also now includes a marine carbon cycle (HadOCC) (Williams et al., 2013) and an oceanic oxygen cycle (Williams et al., 2014), allowing direct comparisons to biogeochemical cycles.

Currently a very high-resolution version of ~~HadAM3H~~ HadAM3BH is finalising development in Bristol. This uses a resolution of 0.625°×0.4166° (576×433 grid points, N288) as this has been suggested as a minimum resolution for realistic simulation of the hydrological cycle (Demory et al., 2014). The model appears to be significantly computationally more efficient (approximately 10× faster) than a similar resolution version of the more recent UK Met Office HadGEM3 model (Walters et al., 2014), because of the lower model top, simplified aerosol physics and major differences in the underlying atmospheric dynamical core.

This paper motivates the continued development and scientific application of the ~~HadCM3~~ HadCM3B family of coupled climate models. Future updates will cover new developments to the presented model version, bug corrections and enhancements.

7 Code Availability

The UK ~~Meteorological~~ Met Office made available the source code of HadCM3 via the Ported Unified Model release (http://www.metoffice.gov.uk/research/collaboration/um-partnership). Enquiries regarding the use of HadCM3 should be directed in the first instance to the UM Partnership team, who can be contacted at UM_collaboration@metoffice.gov.uk.

The main repository for the Met Office Unified Model (UM) ~~at the~~ version corresponding to the model presented here can be ~~found~~ viewed at http://cms.ncas.ac.uk/code_browsers/UM4.5/UMbrowser/index.html.

The code detailing the advances described in this paper is completely contained within the files available as a Supplement to this paper. These files are known as code modification files or "mod" files and should be applied to the original code of the model. This is protected under UK Crown Copyright, as is the base code linked above.

The UM basis files for the simulations described in this paper can be found on the puma.nerc.ac.uk facility (please contact Andy Heaps for access, andy.heaps@ncas.ac.uk). The simulation names are:

- tcsyf: ~~HadCM3-M1~~ HadCM3B-M1
- tcywd: ~~HadCM3-M2~~HadCM3B-M2.1N
- tcyxc: ~~HadCM3-M2~~HadCM3B-M2.2N
- tdbad: ~~HadCM3L-M2~~HadCM3BL-M2.1N
- tdekd: ~~HadAM3-M2~~HadAM3B-M2.1N
- tdewb: ~~HadAM3H-M2~~HadAM3BH-M2.1N

– tdexb: FAMOUS-M1

– tdeyb: FAMOUS-M2.2N

– tdkym: ~~HadCM3-M2~~HadCM3B-M2.1D

– tdkyn: ~~HadCM3L-M2~~HadCM3BL-M2.1D

– tdkyo: ~~HadAM3-M2~~HadAM3B-M2.1D

8 Data Availability

– The CERES data were obtained from the NASA Langley Research Center CERES ordering tool at (http://ceres.larc.nasa.gov/).

– We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modelling groups for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

– We acknowledge the MODIS/TERRA project for the NPP data, acquired from: http://neo.sci.gsfc.nasa.gov/view.php?datasetId=MOD17A2_M_PSN&year=2001.

– Data from the RAPID-WATCH MOC monitoring project are funded by the Natural Environment Research Council and are freely available from: http://www.rapid.ac.uk/rapidmoc

Acknowledgements.

– The efforts of all of the developers of HadCM3 at the Hadley Centre and UK Met Office is gratefully acknowledged. Without their hard and successful work we would not be able to tackle some important scientific problems.

– Climate simulations were carried out using the computational facilities of the Advanced Computing Research Centre, University of Bristol (http://www.bris.ac.uk/acrc) (Bluecrystal).

– E. Armstrong was funded by NERC (NE/L501554/1).

– M. P. S. Badger was funded by NERC (NE/J008591/1).

– T. Davies-Barnard was funded by the European Commission's Seventh Framework Program grant agreement 282672 (EMBRACE) and EU grant ERC-2013-CoG-617313 (PaleoGenie).

– A. Farnsworth was funded by NERC (NE/K014757/1; NE/I005722/1; NE/I005714/1)

– P. O. Hopcroft was funded by NERC (NE/I010912/1).

– A. T. Kennedy was funded by NERC (NE/L002434/1).

– N. S. Lord was funded by RWM Limited via a framework contract with Amec Foster Wheeler, who were supported by Quintessa.

– A. Marzocchi was funded by the National Science Foundation (NSF) award #1536454

– W. H. G. Roberts was funded by the Leverhulme Trust.

– G. J. L. Tourte was funded through the advanced ERC grant “The Greenhouse Earth System” (T-GRES, project reference 340923), awarded to R. D. Pancost.

References

Adam, J. C., Clark, E. A., Lettenmaier, D. P., and Wood, E. F.: Correction of Global Precipitation Products for Orographic Effects, Journal of Climate, 19, 15–38, doi:10.1175/JCLI3604.1, 2006.

Antonov, J. I., Seidov, D., Boyer, T. P., Locarnini, R. A., Mishonov, A. V., Garcia, H. E., Baranova, O. K., Zweng, M. M., and Johnson, D. R.: Volume 2: Salinity, World Ocean Atlas 2009, 2010.

Arakawa, A. and Lamb, V. R.: Computational Design of the Basic Dynamical Processes of the UCLA General Circulation Model, in: General Circulation Models of the Atmosphere, edited by Chang, J., vol. 17 of *Methods in Computational Physics: Advances in Research and Applications*, pp. 173–265, Elsevier, doi:10.1016/B978-0-12-460817-7.50009-4, 1977.

Armstrong, E., Valdes, P., House, J., and Singarayer, J.: The Role of CO2 and Dynamic Vegetation on the Impact of Temperate Land-Use Change in the HadCM3 Coupled Climate Model, Earth Interactions, 20, 1–20, doi:10.1175/EI-D-15-0036.1, 2016.

Arnell, N. W., Hudson, D. A., and Jones, R. G.: Climate change scenarios from a regional climate model: Estimating change in runoff in southern Africa, Journal of Geophysical Research: Atmospheres, 108, doi:10.1029/2002JD002782, 4519, 2003.

Beerling, D. J., Fox, A., Stevenson, D. S., and Valdes, P. J.: Enhanced chemistry-climate feedbacks in past greenhouse worlds, Proc Natl Acad Sci USA, 108, 9770–9774, doi:10.1073/pnas.1102409108, 2011.

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Ménard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B., and Harding, R. J.: The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes, Geoscientific Model Development, 4, 677–699, doi:10.5194/gmd-4-677-2011, 2011.

Betts, R. A., Cox, P. M., Collins, M., Harris, P. P., Huntingford, C., and Jones, C. D.: The role of ecosystem-atmosphere interactions in simulated Amazonian precipitation decrease and forest dieback under global climate warming, Theoretical and applied climatology, 78, 157–175, 2004.

Bhaskaran, B., Jones, R. G., Murphy, J. M., and Noguer, M.: Simulations of the Indian summer monsoon using a nested regional climate model: domain size experiments, Climate Dynamics, 12, 573–587, doi:10.1007/BF00216267, 1996.

Booth, B. B. B., Jones, C. D., Collins, M., Totterdell, I. J., Cox, P. M., Sitch, S., Huntingford, C., Betts, R. A., Harris, G. R., and Lloyd, J.: High sensitivity of future global warming to land carbon cycle processes, Environmental Research Letters, 7, 024 002, doi:10.1088/1748-9326/7/2/024002, 2012.

Braconnot, P., Otto-Bliesner, B., Harrison, S., Joussaume, S., Peterchmitt, J.-Y., Abe-Ouchi, A., Crucifix, M., Driesschaert, E., Fichefet, T., Hewitt, C. D., Kageyama, M., Kitoh, A., Laîné, A., Loutre, M.-F., Marti, O., Merkel, U., Ramstein, G., Valdes, P., Weber, S. L., Yu, Y., and Zhao, Y.: Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum – Part 1: experiments and large-scale features, Climate of the Past, 3, 261–277, doi:10.5194/cp-3-261-2007, 2007.

Bradshaw, C. D., Lunt, D. J., Flecker, R., and Davies-Barnard, T.: Disentangling the roles of late Miocene palaeogeography and vegetation - Implications for climate sensitivity, Palaeogeography, Palaeoclimatology, Palaeoecology, 417, 17–34, 2015.

Bragg, F. J., Lunt, D. J., and Haywood, A. M.: Mid-Pliocene climate modelled using the UK Hadley Centre Model: PlioMIP Experiments 1 and 2, Geosci. Model Dev., 5, 1109–1125, doi:10.5194/gmd-5-1109-2012, gMD, 2012.

Brooks, R. H. and Corey, A. T.: Hydraulic properties of porous media, Hydrology Papers, 3, 1964.

Bryan, K. and Cox, M. D.: An Approximate Equation of State for Numerical Models of Ocean Circulation, Journal of Physical Oceanography, 2, 510–514, doi:10.1175/1520-0485(1972)002<0510:AAEOSF>2.0.CO;2, 1972.

Bushell, A.: Unified Model User Guide, chap. Clouds, p. 27, in: Matthews (1998), http://www.ukscience.org/_Media/UM_User_Guide.pdf, 1998.

Clark, D. B., Mercado, L. M., Sitch, S., Jones, C. D., Gedney, N., Best, M. J., Pryor, M., Rooney, G. G., Essery, R. L. H., Blyth, E., Boucher, O., Harding, R. J., Huntingford, C., and Cox, P. M.: The Joint UK Land Environment Simulator (JULES), model description – Part 2: Carbon fluxes and vegetation dynamics, Geoscientific Model Development, 4, 701–722, doi:10.5194/gmd-4-701-2011, 2011.

Claussen, M., Mysak, L., Weaver, A., Crucifix, M., Fichefet, T., Loutre, M.-F., Weber, S., Alcamo, J., Alexeev, V., Berger, A., Calov, R., Ganopolski, A., Goosse, H., Lohmann, G., Lunkeit, F., Mokhov, I., Petoukhov, V., Stone, P., and Wang, Z.: Earth system models of intermediate complexity: closing the gap in the spectrum of climate system models, Climate Dynamics, 18, 579–586, doi:10.1007/s00382-001-0200-1, 2002.

Collins, M., Tett, S. F. B., and Cooper, C.: The internal climate variability of HadCM3, a version of the Hadley Centre coupled model without flux adjustments, Climate Dynamics, 17, 61–81, doi:10.1007/s003820000094, 2001.

Collins, W. J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P., Hinton, T., Hughes, J., Jones, C. D., Joshi, M., Liddicoat, S., Martin, G., O'Connor, F., Rae, J., Senior, C., Sitch, S., Totterdell, I., Wiltshire, A., and Woodward, S.: Development and evaluation of an Earth-System model – HadGEM2, Geoscientific Model Development, 4, 1051–1075, doi:10.5194/gmd-4-1051-2011, 2011.

Cox, P. M.: A primitive equation, 3-dimensional model of the ocean, GFDL Ocean Group Technical Report No. 1. Available from Geophysical Fluid Dynamics Laboratory, P.O. Box 308, Princeton, New Jersey, 08542, 1984.

Cox, P. M.: Description of the TRIFFID dynamic global vegetation model: Hadley Centre Technical Note 24, Tech. rep., Met Office Hadley Centre, Exeter, UK, http://www.metoffice.gov.uk/media/pdf/9/h/HCTN_24.pdf, 2001.

Cox, P. M., Huntingford, C., and Harding, R. J.: A canopy conductance and photosynthesis model for use in a GCM land surface scheme, Journal of Hydrology, 212–213, 79–94, doi:10.1016/S0022-1694(98)00203-0, 1998.

Cox, P. M., Betts, R. A., Bunton, C. B., Essery, R. L. H., Rowntree, P. R., and Smith, J.: The impact of new land surface physics on the GCM simulation of climate and climate sensitivity, Climate Dynamics, 15, 183–203, doi:10.1007/s003820050276, 1999.

Cox, P. M., Betts, R. A., Jones, C. D., Spall, S. A., and Totterdell, I. J.: Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model, Nature, 408, 184–187, doi:10.1038/35041539, 2000.

Cusack, S., Slingo, A., Edwards, J. M., and Wild, M.: The radiative impact of a simple aerosol climatology on the Hadley Centre atmospheric GCM, Quarterly Journal of the Royal Meteorological Society, 124, 2517–2526, doi:10.1002/qj.49712455117, 1998.

Davies-Barnard, T., Valdes, P. J., Jones, C. D., and Singarayer, J. S.: Sensitivity of a coupled climate model to canopy interception capacity, Climate Dynamics, 42, 1715–1732, doi:10.1007/s00382-014-2100-1, 2014.

Day, J. J., Bamber, J., Valdes, P. J., and Kohler, J.: The impact of a seasonally ice free Arctic Ocean on the temperature, precipitation and surface mass balance of Svalbard, The Cryosphere, 6, 35–50, doi:10.5194/tc-6-35-2012, 2012.

Day, J. J., Bamber, J. L., and Valdes, P. J.: The Greenland Ice Sheet's surface mass balance in a seasonally sea ice-free Arctic, Journal of Geophysical Research: Earth Surface, 118, 1533–1544, doi:10.1002/jgrf.20112, 2013.

Demory, M.-E., Vidale, P. L., Roberts, M. J., et al.: The role of horizontal resolution in simulating drivers of the global hydrological cycle, Clim Dyn, 42, 2201–2225, doi:10.1007/s00382-013-1924-4, 2014.

Dolan, A. M., Haywood, A. M., Hunter, S. J., Tindall, J. C., Dowsett, H. J., Hill, D. J., and Pickering, S. J.: Modelling the enigmatic Late Pliocene Glacial Event – Marine Isotope Stage M2, Global and Planetary Change, 128, 47–60, doi:10.1016/j.gloplacha.2015.02.001, 2015.

Durman, C. F., Gregory, J. M., Hassell, D. C., Jones, R. G., and Murphy, J. M.: A comparison of extreme European daily precipitation simulated by a global and a regional climate model for present and future climates, Quarterly Journal of the Royal Meteorological Society, 127, 1005–1015, doi:10.1002/qj.49712757316, 2001.

Edwards, J.: ‘Radiation’. Unified Model User Guide. Version 4.4., Report, The Meteorological Office, UK, Bracknell, Berkshire, UK, http://www.ukscience.org/_Media/UM_User_Guide.pdf, 1998.

Edwards, J. M. and Slingo, A.: Studies with a flexible new radiation code. I: Choosing a configuration for a large-scale model, Quarterly Journal of the Royal Meteorological Society, 122, 689–719, doi:10.1002/qj.49712253107, 1996.

Edwards, M. O.: Global gridded elevation and bathymetry on 5-minute geographic grid (ETOPO5), NOAA, National Geophysical Data Center, 1989.

Essery, R. L. H., Best, M. J., and Cox, P. M.: MOSES 2.2 Technical Documentation, Tech. rep., Hadley Centre, Met Office, http://www.metoffice.gov.uk/media/pdf/9/j/HCTN_30.pdf, hadley Centre technical note 30, 2001.

Essery, R. L. H., Best, M. J., Betts, R. A., Cox, P. M., and Taylor, C. M.: Explicit Representation of Subgrid Heterogeneity in a GCM Land Surface Scheme, Journal of Hydrometeorology, 4, 530–543, doi:10.1175/1525-7541(2003)004<0530:EROSHI>2.0.CO;2, 2003.

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, Geoscientific Model Development, 9, 1937–1958, doi:10.5194/gmd-9-1937-2016, 2016a.

Eyring, V., Righi, M., Lauer, A., Evaldsson, M., Wenzel, S., Jones, C., Anav, A., Andrews, O., Cionni, I., Davin, E. L., Deser, C., Ehbrecht, C., Friedlingstein, P., Gleckler, P., Gottschaldt, K.-D., Hagemann, S., Juckes, M., Kindermann, S., Krasting, J., Kunert, D., Levine, R., Loew, A., Mäkelä, J., Martin, G., Mason, E., Phillips, A. S., Read, S., Rio, C., Roehrig, R., Senftleben, D., Sterl, A., van Ulft, L. H., Walton, J., Wang, S., and Williams, K. D.: ESMValTool (v1.0) – a community diagnostic and performance metrics tool for routine evaluation of Earth system models in CMIP, Geoscientific Model Development, 9, 1747–1802, doi:10.5194/gmd-9-1747-2016, 2016b.

Ferraro, R., Waliser, D. E., Glecker, P., Taylor, K. E., and Eyring, V.: Evolving obs4MIPS to support the Sixth Coupled Model Intercomparison Project (CMIP6), Am Metereol. Soc., pp. 131–133, doi:10.1175/BAMS-D-14-00216.1, 2015.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S., Collins, W., Cox, P. M., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., and Rummukainen, M.: Evaluation of Climate Models, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P., chap. 9, pp. 741–866, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, doi:10.1017/CBO9781107415324.020, 2013.

Foreman, S. J.: Unified Model Documentaiton Paper Number 40, The Ocean Model, Report, The Met. Office, http://cms.ncas.ac.uk/documents/vn4.5/p040.pdf, 2005.

Gent, P. R. and McWilliams, J. C.: Isopycnal Mixing in Ocean Circulation Models, Journal of Physical Oceanography, 20, 150–155, 1990.

Gleckler, P. J., Taylor, K. E., and Doutriaux, C.: Performance metrics for climate models, Journal of Geophysical Research: Atmospheres, 113, doi:10.1029/2007JD008972, d06104, 2008.

Goosse, H., Brovkin, V., Fichefet, T., Haarsma, R., Huybrechts, P., Jongma, J., Mouchet, A., Selten, F., Barriat, P.-Y., Campin, J.-M., Deleersnijder, E., Driesschaert, E., Goelzer, H., Janssens, I., Loutre, M.-F., Morales Maqueda, M. A., Opsteegh, T., Mathieu, P.-P., Munhoven, G., Pettersson, E. J., Renssen, H., Roche, D. M., Schaeffer, M., Tartinville, B., Timmermann, A., and Weber, S. L.: Description of the Earth system model of intermediate complexity LOVECLIM version 1.2, Geoscientific Model Development, 3, 603–633, doi:10.5194/gmd-3-603-2010, 2010.

Gordon, C., Cooper, C., Senior, C. A., Banks, H., Gregory, J. M., Johns, T. C., Mitchell, J. F. B., and Wood, R. A.: The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments, Climate Dynamics, 16, 147–168, doi:10.1007/s003820050010, 2000.

Grant, A.: Unified Model User Guide, chap. Convection, p. 32, in: Matthews (1998), http://www.ukscience.org/_Media/UM_User_Guide.pdf, 1998.

Gregory, D., Kershaw, R., and Inness, P. M.: Parametrization of momentum transport by convection. II: Tests in single-column and general circulation models, Quarterly Journal of the Royal Meteorological Society, 123, 1153–1183, doi:10.1002/qj.49712354103, 1997.

Gregory, D., Shutts, G. J., and Mitchell, J. R.: A new gravity-wave-drag scheme incorporating anisotropic orography and low-level wave breaking: Impact upon the climate of the UK Meteorological Office Unified Model, Quarterly Journal of the Royal Meteorological Society, 124, 463–493, doi:10.1002/qj.49712454606, 1998.

Gregory, J. M., Ingram, W. J., Palmer, M. A., Jones, G. S., Stott, P. A., Thorpe, R. B., Lowe, J. A., Johns, T. C., and Williams, K. D.: A new method for diagnosing radiative forcing and climate sensitivity, Geophysical Research Letters, 31, doi:Artn L03205 10.1029/2003gl018747, 2004.

Gregory, J. M., Browne, O. J. H., Payne, A. J., Ridley, J. K., and Rutt, I. C.: Modelling large-scale ice-sheet–climate interactions following glacial inception, Clim Past, 8, 1565–1580, doi:10.5194/cp-8-1565-2012, 2012.

Griffies, S., Gnanadesikan, A., Pacanowski, R., et al.: Oceanic isopycnal mixing by coordinate rotation, J Phys Oceanogr, 12, 1154–1158, 1982.

Griffies, S., Gnanadesikan, A., Pacanowski, R., et al.: Isoneutral diffusion in a z-coordinate ocean model, J Phys Oceanogr, 28, 805–830, 1998.

HadGEM2 Development Team: The HadGEM2 family of Met Office Unified Model climate configurations, Geosci Model Dev, 4, 723–757, doi:10.5194/gmd-4-723-2011, 2011.

Hawkins, E., Smith, R. S., Allison, L. C., Gregory, J. M., Woollings, T. J., Pohlmann, H., and de Cuevas, B.: Bistability of the Atlantic overturning circulation in a global climate model and links to ocean freshwater transport, Geophysical Research Letters, 38, doi:Artn L10605 10.1029/2011gl047208, 2011.

Haywood, A. M., Valdes, P. J., and Markwick, P. J.: Cretaceous (Wealden) climates: a modelling perspective, Cretaceous Research, 25, 303–311, doi:10.1016/j.cretres.2004.01.005, 2004.

Hewitt, C. D., Broccoli, A. J., Mitchell, J. F. B., and Stouffer, R. J.: A coupled model study of the Last Glacial Maximum: Was part of the North Atlantic relatively warm?, Geophysical Research Letters, 28, 1571–1574, doi:10.1029/2000GL012575, 2001.

Hibler, W. D.: A Dynamic Thermodynamic Sea Ice Model, Journal of Physical Oceanography, 9, 815–846, doi:10.1175/1520-0485(1979)009<0815:ADTSIM>2.0.CO;2, 1979.

Hopcroft, P., Valdes, P., Wania, R., and Beerling, D.: Limited response of peatland CH4 emissions to abrupt Atlantic Ocean circulation changes in glacial climates, Climate of the Past, 10, 137–154, doi:10.5194/cp-10-137-2014, 2014.

Hudson, D. A. and Jones, R. G.: Regional Climate Models Simulations of Present-Day and Future Climate over Southern Africa, Tech. rep., Met Office Hadley Center, Exeter, UK, https://digital.nmla.metoffice.gov.uk/download/file/sdb%3AdigitalFile%7Ca55db57e-b268-49ba-9cc3-d48553b244f5/, 2002.

Hughes, J. K., Valdes, P. J., and Betts, R. A.: Dynamics of a global-scale vegetation model, Ecological Modelling, 198, 452–462, doi:10.1016/j.ecolmodel.2006.05.020, 2006.

Ingram, W. S., Woodward, S., and Edwards, J.: Unified Model Documentation Paper: Radiation, Report 23, The Meteorological Office, UK, 1997.

Irvine, P. J., Ridgwell, A., and Lunt, D. J.: Assessing the regional disparities in geoengineering impacts, Geophysical Research Letters, 37, L18 702, doi:10.1029/2010GL044447, 2010.

Jackson, L. and Vellinga, M.: Multidecadal to Centennial Variability of the AMOC: HadCM3 and a Perturbed Physics Ensemble, Journal of Climate, 26, 2390–2407, doi:10.1175/JCLI-D-11-00601.1, 2013.

Johns, T. C., Carnell, R. E., Crossley, J. F., Gregory, J. M., Mitchell, J. F. B., Senior, C. A., Tett, S. F. B., and Wood, R. A.: The second Hadley Centre coupled ocean-atmosphere GCM: model description, spinup and validation, Climate Dynamics, 13, 103–134, doi:10.1007/s003820050155, 1997.

Johns, T. C., Gregory, J. M., Ingram, W. J., Johnson, C. E., Lowe, J. A., Mitchell, J. F. B., Roberts, D. L., Sexton, D. M. H., Stevenson, D., Tett, S. F. B., and Woodage, M. J.: Anthropogenic climate change for 1860 to 2100 simulated with the HadCM3 model under updated emissions scenarios, Climate Dynamics, 20, 583–612, doi:10.1007/s00382-002-0296-y, 2003.

Jones, C.: A Fast Ocean GCM without Flux Adjustments, Journal of Atmospheric and Oceanic Technology, 20, 1857–1868, doi:10.1175/1520-0426(2003)020<1857:AFOGWF>2.0.CO;2, 2003.

Jones, C., Gregory, J., Thorpe, R., Cox, P., Murphy, J., Sexton, D., and Valdes, P. J.: Systematic optimisation and climate simulation of FAMOUS, a fast version of HadCM3, Climate Dynamics, 25, 189–204, doi:10.1007/s00382-005-0027-2, 2005.

Jones, R. G., Murphy, J. M., and Noguer, M.: Simulation of climate change over europe using a nested regional-climate model. I: Assessment of control climate, including sensitivity to location of lateral boundaries, Quarterly Journal of the Royal Meteorological Society, 121, 1413–1449, doi:10.1002/qj.49712152610, 1995.

Kennedy, A. T., Farnsworth, A., Lunt, D. J., Lear, C. H., and Markwick, P. J.: Atmospheric and oceanic impacts of Antarctic glaciation across the Eocene–Oligocene transition, Philosophical Transactions of the Royal Society, A, 373, doi:10.1098/rsta.2014.0419, 2015.

Kraus, E. B. and Turner, J. S.: A one-dimensional model of the seasonal thermocline II. The general theory and its consequences, Tellus, 19, 98–106, doi:10.1111/j.2153-3490.1967.tb01462.x, 1967.

Large, W. G., McWilliams, J. C., and Doney, S. C.: Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization, Reviews of Geophysics, 32, 363–403, doi:10.1029/94RG01872, 1994.

Legg, S., Briegleb, B., Chang, Y., Chassignet, E. P., Danabasoglu, G., Ezer, T., Gordon, A. L., Griffies, S., Hallberg, R., Jackson, L., Large, W., Ozgokmen, T. M., Peters, H., Price, J., Riemenschneider, U., Wu, W. L., Xu, X. B., and Yang, J. Y.: Improving Oceanic Overflow Representation in Climate Models: The Gravity Current Entrainment Climate Process Team, Bulletin of the American Meteorological Society, 90, 657–670, doi:10.1175/2008bams2667.1, 2009.

Liu, W., Liu, Z. Y., and Brady, E. C.: Why is the AMOC Monostable in Coupled General Circulation Models?, Journal of Climate, 27, 2427–2443, doi:10.1175/Jcli-D-13-00264.1, 2014.

Locarnini, R. A., Mishonov, A. V., Antonov, J. I., Boyer, T. P., Garcia, H. E., Baranova, O. K., Zweng, M. M., , and Johnson, D. R.: Volume 1: Temperature, World Ocean Atlas 2009, 2010.

Loeb, N., Wang, H., Cheng, A., Kato, S., Fasullo, J., X, K.-M., and Allan, R.: Observational constraints on atmospheric and oceanic cross-equatorial heat transports: revisiting the precipitation asymmetry problem in climate models, Clim Dyn, 46, 3239–3257, doi:10.1007/s00382-015-2766-z, 2016.

Loptson, C. A., Lunt, D. J., and Francis, J. E.: Investigating vegetation–climate feedbacks during the early Eocene, Climate of the Past, 10, 419–436, doi:10.5194/cp-10-419-2014, 2014.

Loveland, T. R. and Belward, A. S.: The IGBP-DIS global 1km land cover data set, DISCover: First results, International Journal of Remote Sensing, 18, 3289–3295, doi:10.1080/014311697217099, 1997.

Lunt, D. J., Ridgwell, A., Valdes, P. J., and Seale, A.: “Sunshade World”: A fully coupled GCM evaluation of the climatic impacts of geoengineering, Geophysical Research Letters, 35, L12 710, doi:10.1029/2008GL033674, 2008.

Lunt, D. J., Farnsworth, A., Loptson, C., Foster, G. L., Markwick, P., O’Brien, C. L., Pancost, R. D., Robinson, S. A., and Wrobel, N.: Palaeogeographic controls on climate and proxy interpretation, Clim. Past, 12, 1181–1198, doi:10.5194/cp-12-1181-2016, cP, 2016.

Malhi, Y., Aragão, L. E. O. C., Galbraith, D., Huntingford, C., Fisher, R., Zelazowski, P., Sitch, S., McSweeney, C., and Meir, P.: Exploring the likelihood and mechanism of a climate-change-induced dieback of the Amazon rainforest, Proceedings of the National Academy of Sciences, 106, 20 610–20 615, doi:10.1073/pnas.0804619106, 2009.

Marzocchi, A., Hirschi, J. J.-M., Holliday, N. P., Cunningham, S. A., Blaker, A. T., and Coward, A. C.: The North Atlantic subpolar circulation in an eddy-resolving global ocean model, Journal of Marine Systems, 142, 126–143, doi:10.1016/j.jmarsys.2014.10.007, 2015a.

Marzocchi, A., Lunt, D. J., Flecker, R., Bradshaw, C. D., Farnsworth, A., and Hilgen, F. J.: Orbital control on late Miocene climate and the North African monsoon: insight from an ensemble of sub-precessional simulations, Climate of the Past, 11, 1271–1295, 2015b.

Matthews, D., ed.: Unified Model User Guide, The Meteorological Office, Bracknell, Berkshire, UK, http://www.ukscience.org/_Media/UM_User_Guide.pdf, 1998.

McCarthy, G., Frajka-Williams, E., Johns, W. E., Baringer, M. O., Meinen, C. S., Bryden, H. L., Rayner, D., Duchez, A., Roberts, C., and Cunningham, S. A.: Observed interannual variability of the Atlantic meridional overturning circulation at 26.5 degrees N, Geophysical Research Letters, 39, doi:1029/2012gl052933, 2012.

Msadek, R., Johns, W. E., Yeager, S. G., Danabasoglu, G., Delworth, T. L., and Rosati, A.: The Atlantic Meridional Heat Transport at 26.5 degrees N and Its Relationship with the MOC in the RAPID Array and the GFDL and NCAR Coupled Models, Journal of Climate, 26, 4335–4356, doi:10.1175/Jcli-D-12-00081.1, 2013.

Murphy, J. M., Booth, B. B. B., Collins, M., Harris, G. R., Sexton, D. M. H., and Webb, M. J.: A methodology for probabilistic predictions of regional climate change from perturbed physics ensembles, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 365, 1993–2028, doi:10.1098/rsta.2007.2077, 2007.

New, M., Lister, D., Hulme, M., and Makin, I.: A high-resolution data set of surface climate over global land areas, Climate Research, 21, 1–25, doi:10.3354/cr021001, 2002.

Pacanowski, R. and Philander, S.: Parametrisation of vertical mixing in numerical models of tropical oceans, J Phys Oceanogr, 11, 1443–1451, 1981.

Pardaens, A. K., Banks, H. T., Gregory, J. M., and Rowntree, P. R.: Freshwater transports in HadCM3, Climate Dynamics, 21, 177–195, doi:10.1007/s00382-003-0324-6, 2003.

Pope, V. D., Gallani, M. L., Rowntree, P. R., and Stratton, R. A.: The impact of new physical parametrizations in the Hadley Centre climate model: HadAM3, Climate Dynamics, 16, 123–146, doi:10.1007/s003820050009, 2000.

Ramankutty, N. and Foley, J. A.: Estimating historical changes in global land cover: Croplands from 1700 to 1992, Global Biogeochemical Cycles, 13, 997–1027, doi:10.1029/1999GB900046, 1999.

Reichler, T. and Kim, J.: How well do coupled models simulate today’s climate?, Bulletin of the American Meteorological Society, 89, 303–311, doi:10.1175/BAMS-89-3-303, http://journals.ametsoc.org/doi/abs/10.1175/BAMS-89-3-303, 2008.

Rickard, G.: Unified Model Documentation No. 59: Ocean Models and the Implementation of Vertical Diffusion and Vertical Mixing, Report, The Met. Office, http://cms.ncas.ac.uk/documents/vn4.5/p059.pdf, 1999.

Ridgwell, A., Singarayer, J. S., Hetherington, A. M., and Valdes, P. J.: Tackling Regional Climate Change By Leaf Albedo Bio-geoengineering, Current Biology, 19, 146–150, doi:10.1016/j.cub.2008.12.025, 2009.

Ridley, J. K., Huybrechts, P., Gregory, J. M., and Lowe, J. A.: Elimination of the Greenland ice sheet in a high CO2 climate, Journal of Climate, 18, 3409–3427, doi:Doi 10.1175/Jcli3482.1, 2005.

Roberts, C. D., Garry, F. K., and Jackson, L. C.: A Multimodel Study of Sea Surface Temperature and Subsurface Density Fingerprints of the Atlantic Meridional Overturning Circulation, Journal of Climate, 26, 9155–9174, doi:10.1175/Jcli-D-12-00762.1, 2013.

Roberts, W. H. G. and Valdes, P. J.: Green Mountains and White Plains: the effect of Northern Hemisphere ice sheets on the global energy budget., J. Climate., p. in press, doi:10.1175/JCLI-D-15-0846.1, 2017.

Roberts, W. H. G., Valdes, P. J., and Singarayer, J. S.: Can energy fluxes be used to interpret glacial precipitation changes in the Tropics?, Geophysical Research Letters, in review.

Roberts, W. H. G., Valdes, P. J., and Payne, A. J.: Topography's crucial role in Heinrich Events, Proc Natl Acad Sci USA, 111, 16 688–16 693, doi:10.1073/pnas.1414882111, 2014.

Sarojini, B. B., Gregory, J. M., Tailleux, R., Bigg, G. R., Blaker, A. T., Cameron, D. R., Edwards, N. R., Megann, A. P., Shaffrey, L. C., and Sinha, B.: High frequency variability of the Atlantic meridional overturning circulation, Ocean Science, 7, 471–486, doi:10.5194/os-7-471-2011, 2011.

Schaller, N., Kay, A. L., Lamb, R., Massey, N. R., van Oldenborgh, G. J., Otto, F. E. L., Sparrow, S. N., Vautard, R., Yiou, P., Ashpole, I., Bowery, A., Crooks, S. M., Haustein, K., Huntingford, C., Ingram, W. J., Jones, R. G., Legg, T., Miller, J., Skeggs, J., Wallom, D., Weisheimer, A., Wilson, S., Stott, P. A., and Allen, M. R.: Human influence on climate in the 2014 southern England winter floods and their impacts, Nature Clim. Change, 6, 627–634, doi:10.1038/nclimate2927, article, 2016.

Sijp, W. P.: Characterising meridional overturning bistability using a minimal set of state variables, Climate Dynamics, 39, 2127–2142, doi:10.1007/s00382-011-1249-0, 2012.

Simmons, A. J. and Strüfing, R.: Numerical forecasts of stratospheric warming events using a model with a hybrid vertical coordinate, Quarterly Journal of the Royal Meteorological Society, 109, 81–111, doi:10.1002/qj.49710945905, 1983.

Singarayer, J. S. and Valdes, P. J.: High-latitude climate sensitivity to ice-sheet forcing over the last 120kyr, Quaternary Science Reviews, 29, 43–55, doi:10.1016/j.quascirev.2009.10.011, 2010.

Singarayer, J. S., Ridgwell, A., and Irvine, P.: Assessing the benefits of crop albedo bio-geoengineering, Environmental Research Letters, 4, 045 110, http://stacks.iop.org/1748-9326/4/i=4/a=045110, 2009.

Sitch, S., Huntingford, C., Gedney, N., Levy, P. E., Lomas, M., Piao, S. L., Betts, R., Ciais, P., Cox, P. M., Friedlingstein, P., Jones, C. D., Prentice, I. C., and Woodward, F. I.: Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs), Global Change Biology, 14, 2015–2039, doi:10.1111/j.1365-2486.2008.01626.x, 2008.

Smeed, D., McCarthy, G., Rayner, D., Moat, B. I., Johns, W. E., Baringer, M., and Meinen, C. S.: Atlantic meridional overturning circulation observed by the RAPID-MOCHA-WBTS (RAPID-Meridional Overturning Circulation and Heatflux Array-Western Boundary Time Series) array at 26N from 2004 to 2014., British Oceanographic Data Centre - Natural Environment Research Council, UK, doi:10/6qb, 2015.

Smith, D. M., Cusack, S., Colman, A. W., Folland, C. K., Harris, G. R., and Murphy, J. M.: Improved Surface Temperature Prediction for the Coming Decade from a Global Climate Model, Science, 317, 796–799, doi:10.1126/science.1139540, 2007.

Smith, D. M., Eade, R., Dunstone, N. J., Fereday, D., Murphy, J. M., Pohlmann, H., and Scaife, A. A.: Skilful multi-year predictions of Atlantic hurricane frequency, Nature Geosci, 3, 846–849, doi:10.1038/ngeo1004, 2010.

Smith, R. S.: The FAMOUS climate model (versions XFXWB and XFHCC): description update to version XDBUA, Geoscientific Model Development, 5, 269–276, doi:10.5194/gmd-5-269-2012, 2012.

Smith, R. S., Gregory, J. M., and Osprey, A.: A description of the FAMOUS (version XDBUA) climate model and control run, Geoscientific Model Development, 1, 53–68, doi:10.5194/gmd-1-53-2008, 2008.

Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Avery, K., Tignor, M., and Miller, H., eds.: IPCC, 2007: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2007.

Spencer, H., Sutton, R., and Slingo, J. M.: El Niño in a Coupled Climate Model: Sensitivity to Changes in Mean State Induced by Heat Flux and Wind Stress Corrections, Journal of Climate, 20, 2273–2298, doi:10.1175/JCLI4111.1, 2007.

Stainforth, D. A., Aina, T., Christensen, C., Collins, M., Faull, N., Frame, D. J., Kettleborough, J. A., Knight, S., Martin, A., Murphy, J. M., Piani, C., Sexton, D., Smith, L. A., Spicer, R. A., Thorpe, A. J., and Allen, M. R.: Uncertainty in predictions of the climate response to rising levels of greenhouse gases, Nature, 433, 403–406, doi:10.1038/nature03301, 2005.

Stott, P. A. and Kettleborough, J. A.: Origins and estimates of uncertainty in predictions of twenty-first century temperature rise, Nature, 416, 723–726, 2002.

Stott, P. A., Tett, S. F. B., Jones, G. S., Allen, M. R., Mitchell, J. F. B., and Jenkins, G. J.: External control of 20th century temperature by natural and anthropogenic forcings, Science, 290, 2133–2137, 2000.

Svirezhev, Y. M.: Thermodynamics and ecology, Ecological Modelling, 132, 11–22, doi:10.1016/S0304-3800(00)00301-X, 2000.

Tett, S. F. B., Betts, R., Crowley, T. J., Gregory, J., Johns, T. C., Jones, A., Osborn, T. J., Öström, E., Roberts, D. L., and Woodage, M. J.: The impact of natural and anthropogenic forcings on climate and hydrology since 1550, Climate Dynamics, 28, 3–34, doi:10.1007/s00382-006-0165-1, 2006.

Toniazzo, T., Collins, M., and Brown, J.: The variation of ENSO characteristics associated with atmospheric parameter perturbations in a coupled model, Climate Dynamics, 30, 643–656, doi:10.1007/s00382-007-0313-2, 2007.

Trenberth, K. E. and Fasullo, J. T.: An apparent hiatus in global warming?, Earth's Future, 1, 19–32, doi:10.1002/2013EF000165, 2013.

UNESCO: Tenth report of the joint panel on oceanographic tables and standards., Report, http://unesdoc.unesco.org/images/0004/000461/046148eb.pdf, 1981.

Valdes, P. J., Beerling, D. J., and Johnson, C. E.: The ice age methane budget, Geophys Res Lett, 32, L02 704, doi:10.1029/2004GL021004, 2005.

Van der Wal, A.: Unified Model User Guide, chap. Radiation, p. 11, in: Matthews (1998), http://www.ukscience.org/_Media/UM_User_Guide.pdf, 1998.

Visbeck, M., Marshall, J., Haine, T., and Spall, M.: On the specification of eddy transfer coefficients in coarse resolution ocean circulation models, J Phys Oceanogr, 27, 381–402, 1997.

Walters, D. N. et al.: The Met Office Unified Model Global Atmosphere 4.0 and JULES Global Land 4.0 configurations, Geosci Mod Dev., pp. 361–386, doi:10.5194/gmd-7-361-2014, 2014.

Warrilow, D., Sangster, A., and Slingo: Modelling of land surface processes and their influence on European climate, Met 0ffice 20 (Dynamical Climatology Branch), Meteorological Office, 1986.

Warrilow, D. A. and Buckley, E.: The Impact of Land Surface Processes on the Moisture Budget of a Climate Model, Annales Geophysicae-Atmospheres Hydrospheres and Space Sciences, 7, 439–449, 1989.

Weaver, A. J., Sedlacek, J., Eby, M., Alexander, K., Crespin, E., Fichefet, T., Philippon-Berthier, G., Joos, F., Kawamiya, M., Matsumoto, K., Steinacher, M., Tachiiri, K., Tokos, K., Yoshimori, M., and Zickfeld, K.: Stability of the Atlantic meridional overturning circulation: A model intercomparison, Geophysical Research Letters, 39, doi:10.1029/2012gl053763, 2012.

White, A. A. and Bromley, R. A.: Dynamically consistent, quasi-hydrostatic equations for global models with a complete representation of the Coriolis force, Quarterly Journal of the Royal Meteorological Society, 121, 399–418, doi:10.1002/qj.49712152208, 1995.

Williams, J. H. T., Smith, R. S., Valdes, P. J., Booth, B. B. B., and Osprey, A.: Optimising the FAMOUS climate model: inclusion of global carbon cycling, Geoscientific Model Development, 6, 141–160, doi:10.5194/gmd-6-141-2013, 2013.

Williams, J. H. T., Totterdel, I. J., Halloran, P. R., and Valdes, P. J.: Numerical simulations of oceanic oxygen cycling in the FAMOUS Earth-System model: FAMOUS-ES, version 1.0, Geosci Mod Dev., pp. 1419–1431, doi:10.5194/gmd-7-1419-2014, 2014.

Williams, K. D., Harris, C. M., Bodas-Salcedo, A., et al.: The Met Office Global Coupled model 2.0 (GC2) configuration, Geosci Model Dev, 88, 1509–1524, doi:10.5194/gmd-88-1509-2015, 2015.

Willmott, C. J., Rowe, C. M., and Mintz, Y.: Climatology of the terrestrial seasonal water cycle, Journal of Climatology, 5, 589–606, doi:10.1002/joc.3370050602, 1985.

Wilson, D.: Unified Model User Guide, chap. Precipitation, p. 32, in: Matthews (1998), http://www.ukscience.org/_Media/UM_User_Guide.pdf, 1998.

Wilson, M. F. and Henderson-Sellers, A.: A global archive of land cover and soils data for use in general circulation climate models, Journal of Climatology, 5, 119–143, doi:10.1002/joc.3370050202, 1985.

Zhang, L. P. and Wang, C. Z.: Multidecadal North Atlantic sea surface temperature and Atlantic meridional overturning circulation variability in CMIP5 historical simulations, Journal of Geophysical Research-Oceans, 118, 5772–5791, doi:10.1002/jgrc.20390, 2013.