Reviewer 2

General comments:
This manuscript uses a pattern scaling approach to compare CMIP5 models and observational data related to heterotrophic respiration (RH). In terms of absolute RH, on average the models substantially overestimate global RH; the models also predict a substantial increase in RH over the next century. Some but not all models predict an increase in temporal variability of RH as well. Models show spatial biases and are not particularly well correlated with the dataset. RH correlations with temperature, precipitation, and NPP vary zonally in the models, but not in the same way as the observations. Furthermore, models vary dramatically in their local RH sensitivity to global changes in RH, NPP, and climate variables.

I think it is a useful and important exercise to compare ESM outputs of RH. This analysis is novel and complements previous work on soil carbon stocks and NPP. That said, the paper could benefit from an improved explanation of its goals, expectations, and approaches.

Thanks for the careful reading and useful feedback. We agree that this manuscript needs significant revisions in many areas, but are hopeful that doing so will greatly improve its clarity, methodological rigor, and ultimately impact.

Goals: At the end of the intro, the goal is to determine if pattern scaling can be used to evaluate models. Of course it can. But more importantly, can it yield new insight into the underlying biases or problems with the models? Hopefully so, and the intro could do a better job at identifying the issues that are most likely to be revealed by the technique.

This is a good point. In our revision we will provide more background on pattern scaling, as well as the state of CMIP5 models' carbon cycle performance more generally (Anav et al., 2013; Luo et al., 2016), and using aspects of model behavior to draw inferences about climate- and carbon-cycle response to anthropogenic forcing (e.g. Gillett et al., 2013). In addition, we think that a better discussion of how RH pattern scaling can be treated as a type of emergent constraint (Hoffman et al., 2014; Luo et al., 2015) would be useful.

Expectations: It was very difficult to interpret the myriad results in the paper without a better indication of how to use the results for model diagnosis. This problem could be addressed by including some pseudo-hypotheses about expected patterns in the model comparisons. For instance, if the processes in a model are overly sensitive to temperature, then what pattern in Fig. 4a might be expected? It was really hard to know what a reader should be looking for when viewing the results, especially with so many different models.

This is an excellent suggestion–thank you. In our revision we will clearly lay out specific hypotheses or at least expectations: how, based on our best understanding of the carbon cycle and scaling issues (Bond-Lamberty et al., 2016; Jung et al., 2017; Phillips et al., 2017), models might be expected to behave across latitudinal and global scales.

Approaches: I was not previously familiar with the pattern scaling approach, and although it looks powerful, I had a hard time understanding it from the manuscript text. The methods section on pattern scaling could use some elaboration. In particular, I struggled to understand how a single value could be used as the X with multiple Y values in a regression analysis. I was also unclear about the temporal change component. Was the regression relating the change in Y with the change in X over some time interval? The equation presented in this section needs to be explained more clearly or in more detail (or both).

In our revision we will provide more background on pattern scaling: its strengths and limitations (Tebaldi and Arblaster, 2014), accuracy (Herger et al., 2015; Mitchell, 2003), and perhaps novel statistical approaches to overcome some its limitations (Link et al., 2018). All equation(s) will be more carefully explained.

Specific comments:
I wonder if the authors should consider the implications of non-independence among the models. Previously it has been found that ESMs with the same underlying biogeochemical model have very similar predictions of soil carbon spatial distributions and are not independent. Clearly not all 25 model variants in this paper are independent as some of them generate essentially identical outputs. Is it necessary to show all 25 models? Can they be grouped or aggregated in some sensible way?

We agree, and attempted to do this already (to some degree) in Figure 5. Yes, this is an excellent suggestion and will tie in nicely with the repeated reviewer requests for more clarity in figures and results–we will do much more grouping by ‘model center’ or land model used (e.g. different variants of CESM, etc).

Finally, the manuscript seemed to be missing an overall conclusion about the models and recommendations for future model development. There are a lot of discrepancies with the data and across models identified in the paper. Where should the ESM community be moving with respect to improving predictions of RH? It seemed like some models, like GISS, were outliers, but are there other areas that need attention?
Editorial comments:
Abstract fails to give a set of general conclusions specific to this study analysis.

We agree, and will place more emphasis on synthesizing suggestions for future directions and steps to improve RH modeling.

1:41- Not clear what "This" refers to; here and throughout, specify directly.

Thanks; we will do so.

2:34- I suggest a different formulation of the objective. Almost anything "can" be done. I suspect you were interested in specific aspects of the changes and uncertainty. Can this objective be more informative?

This links well with the suggestion above to have specific hypotheses, or at least expectations. We feel that this will help better define a tight objective, which we agree is currently missing.

2:46- fix reference formatting
5:39- "does not"
6:32- missing "is"
8:24- missing a word in here somewhere
9:33- fix reference formatting

We will fix these issues.

References cited in response

Anav, A., Friedlingstein, P., Kidston, M., Bopp, L., Ciais, P., Cox, P. M., Jones, C. D., Jung, M., Myneni, R. B. and Zhu, Z.: Evaluating the land and ocean components of the global carbon cycle in the CMIP5 earth system models, J. Clim., 26, 6801–6843, 2013.

Bond-Lamberty, B. and Thomson, A. M.: Temperature-associated increases in the global soil respiration record, Nature, 464(7288), 579–582, 2010.

Bond-Lamberty, B., Epron, D., Harden, J. W., Harmon, M. E., Hoffman, F. M., Kumar, J., McGuire, A. D. and Vargas, R.: Estimating heterotrophic respiration at large scales: challenges, approaches, and next steps, Ecosphere, 7(6), d01380, 2016.

Davidson, E. A. and Janssens, I. A.: Temperature sensitivity of soil carbon decomposition and feedbacks to climate change, Nature, 440, 165–173, 2006.

Exbrayat, J.-F., Pitman, A. J., Zhang, Q., Abramowitz, G. and Wang, Y.-P.: Examining soil carbon uncertainty in a global model: response of microbial decomposition to temperature, moisture and nutrient limitation, Biogeosciences, 10, 7095–7108, 2013a.

Exbrayat, J.-F., Pitman, A. J., Abramowitz, G. and Wang, Y.-P.: Sensitivity of net ecosystem exchange and heterotrophic respiration to parameterization uncertainty, J. Geophys. Res., doi:10.1029/2012JD018122, 2013b.

Falloon, P. D., Jones, C. D., Ades, M. and Paul, K. I.: Direct soil moisture controls of future global soil carbon changes: An important source of uncertainty, Global Biochemical Cycles, 25, GB3010, 2011.

Gillett, N. P., Arora, V. K., Matthews, D. and Allen, M. R.: Constraining the Ratio of Global Warming to Cumulative CO2 Emissions Using CMIP5 Simulations, J. Clim., 26(18), 6844–6858, 2013.

Guenet, B., Camino-Serrano, M., Ciais, P., Tifafi, M., Maignan, F., Soong, J. L. and Janssens, I. A.: Impact of priming on global soil carbon stocks, Glob. Chang. Biol., doi:10.1111/gcb.14069, 2018.

Hashimoto, S., Carvalhais, N., Ito, A., Migliavacca, M., Nishina, K. and Reichstein, M.: Global spatiotemporal distribution of soil respiration modeled using a global database, Biogeosciences, 12, 4121–4132, 2015.

Herger, N., Sanderson, B. W. and Knutti, R.: Improved pattern scaling approaches for the use in climate impact studies, Geophys. Res. Lett., in press, doi:10.1002/2015GL063569, 2015.

Hoffman, F. M., Randerson, J. T., Arora, V., Bao, Q., Six, K. D., Cadule, P., Ji, D., Jones, C. D., Kawamiya, M., Khatiwala, S., Lindsay, K., Obata, A., Shevliakova, E., Tjiputra, J., Volodin, E. M. and Wu, T.: Causes and implications of persistent atmospheric carbon dioxide biases in Earth System Models, Journal of Geophysical Research-Biogeosciences, in press, doi:10.1002/2013JG002381, 2014.

Hursh, A., Ballantyne, A., Cooper, L., Maneta, M., Kimball, J. and Watts, J.: The sensitivity of

soil respiration to soil temperature, moisture, and carbon supply at the global scale, Glob. Chang. Biol., 23(5), 2090–2103, 2017.

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W. L., Lasslop, G., Law, B. E., Lindroth, A., Merbold, L., Montagnani, L., Moors, E. J., Papale, D., Sottocornola, M., Vaccari, F. and Williams, C. A.: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, Journal of Geophysical Research-Biogeosciences, 116, G00J07, 2011.

Jung, M., Reichstein, M., Schwalm, C. R., Huntingford, C., Sitch, S., Ahlström, A., Arneth, A., Camps-Valls, G., Ciais, P., Friedlingstein, P., Gans, F., Ichii, K., Jain, A. K., Kato, E., Papale, D., Poulter, B., Raduly, B., Rödenbeck, C., Tramontana, G., Viovy, N., Wang, Y.-P., Weber, U., Zaehle, S. and Zeng, N.: Compensatory water effects link yearly global land CO2 sink changes to temperature, Nature, 541(7638), 516–520, 2017.

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Pongratz, J., Manning, A. C., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Jackson, R. B., Boden, T. A., Tans, P. P., Andrews, O. D., Arora, V. K., Bakker, D. C. E., Barbero, L., Becker, M., Betts, R. A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Cosca, C. E., Cross, J., Currie, K., Gasser, T., Harris, I., Hauck, J., Haverd, V., Houghton, R. A., Hunt, C. W., Hurtt, G., Ilyina, T., Jain, A. K., Kato, E., Kautz, M., Keeling, R. F., Klein Goldewijk, K., Körtzinger, A., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Lima, I., Lombardozzi, D., Metzl, N., Millero, F., Monteiro, P. M. S., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Nojiri, Y., Padín, X. A., Peregon, A., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Reimer, J., Rödenbeck, C., Schwinger, J., Séférian, R., Skjelvan, I., Stocker, B. D., Tian, H., Tilbrook, B., van der Laan-Luijkx, I. T., van der Werf, G. R., van Heuven, S., Viovy, N., Vuichard, N., Walker, A. P., Watson, A. J., Wiltshire, A. J., Zaehle, S. and Zhu, D.: Global Carbon Budget 2017, Earth Syst. Sci. Data Discuss., 1–79, 2017.

Link, R., Bond-Lamberty, B., Hartin, C., Lynch, C. and Kravitz, B.: Computationally efficient emulators for Earth System Models, Geoscientific Model Development, submitted, 2018.

Liu, L., Wang, X., Lajeunesse, M. J., Miao, G., Piao, S., Wan, S., Wu, Y., Wang, Z., Yang, S. and Deng, M.: A cross-biome synthesis of soil respiration and its determinants under simulated precipitation changes, Glob. Chang. Biol., 22(4), 1394–1405, 2016.

Luo, Y., Keenan, T. F. and Smith, M.: Predictability of the terrestrial carbon cycle, Glob. Chang. Biol., 21(5), 1737–1751, 2015.

Luo, Y., Ahlström, A., Allison, S. D., Batjes, N. H., Brovkin, V., Carvalhais, N., Chappell, A., Ciais, P., Davidson, E. A., Finzi, A., Georgiou, K., Guenet, B., Hararuk, O., Harden, J. W., He, Y., Hopkins, F., Jiang, L., Koven, C., Jackson, R. B., Jones, C. D., Lara, M. J., Liang, J., McGuire, A. D., Parton, W., Peng, C., Randerson, J. T., Salazar, A., Sierra, C. A., Smith, M. J., Tian, H., Todd-Brown, K. E. O., Torn, M., van Groenigen, K. J., Wang, Y. P., West, T. O., Wei, Y., Wieder, W. R., Xia, J., Xu, X., Xu, X. and Zhou, T.: Toward more realistic projections of soil carbon dynamics by Earth system models, Global Biogeochem. Cycles, 30(1), 2015GB005239, 2016.

Mahecha, M. D., Reichstein, M., Carvalhais, N., Lasslop, G., Lange, H., Seneviratne, S. I., Vargas, R., Ammann, C., Arain, M. A., Cescatti, A., Janssens, I. A., Migliavacca, M.,

Montagnani, L. and Richardson, A. D.: Global convergence in the temperature sensitivity of respiration at ecosystem level, Science, 329(5993), 838–840, 2010.

McGuire, A. D., Anderson, L. G., Christensen, T. R., Dallimore, S., Guo, L., Hayes, D. J., Heimann, M., Lorenson, T. D., Macdonald, R. W. and Roulet, N. T.: Sensitivity of the carbon cycle in the Arctic to climate change, Ecol. Monogr., 79(4), 523–555, 2009.

Mitchell, T. D.: Pattern Scaling: An Examination of the Accuracy of the Technique for Describing Future Climates, Clim. Change, 60(3), 217–242, 2003.

Moyano, F. E., Manzoni, S. and Chenu, C.: Responses of soil heterotrophic respiration to moisture availability: An exploration of processes and models, Soil Biol. Biochem., 59, 72–85, 2013.

Odum, E. P.: The strategy of ecosystem development, Science, 164, 262–270, 1969.

Phillips, C. L., Bond-Lamberty, B., Desai, A. R., Lavoie, M., Risk, D., Tang, J., Todd-Brown, K. and Vargas, R.: The value of soil respiration measurements for interpreting and modeling terrestrial carbon cycling, Plant Soil, 413(1-2), 1–25, 2017.

Rehfeld, K., Münch, T., Ho, S. L. and Laepple, T.: Global patterns of declining temperature variability from the Last Glacial Maximum to the Holocene, Nature, 554(7692), 356–359, 2018.

Schimel, D., Stephens, B. B. and Fisher, J. B.: Effect of increasing CO2 on the terrestrial carbon cycle, Proceedings of the National Academy of Science, 112(2), 436–441, 2015.

Shao, P., Zeng, X., Moore, D. J. P. and Zeng, X.: Soil microbial respiration from observations and Earth System Models, Environ. Res. Lett., 8(3), 034034, 2013.

Tebaldi, C. and Arblaster, J. M.: Pattern scaling: Its strengths and limitations, and an update on the latest model simulations, Clim. Change, 122(3), 459–471, 2014.

Todd-Brown, K. E. O., Randerson, J. T., Post, W. M., Hoffman, F. M., Tarnocai, C., Schuur, E. A. G. and Allison, S. D.: Causes of variation in soil carbon predictions from CMIP5 Earth system models and comparison with observations, Biogeosciences, 10, 1717–1736, 2013.

Xiao, J., Davis, K. J., Urban, N. M., Keller, K. and Saliendra, N. Z.: Upscaling carbon fluxes from towers to the regional scale: Influence of parameter variability and land cover representation on regional flux estimates, Journal of Geophysical Research-Biogeosciences, 116, GB3027, 2011.