Atmos. Chem. Phys. Discuss., 15, C2381–C2387, 2015
www.atmos-chem-phys-discuss.net/15/C2381/2015/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Improvement of climate predictions and reduction of their uncertainties using learning algorithms" *by* E. Strobach and G. Bel

Anonymous Referee #1

Received and published: 8 May 2015

Strong similarities to past work

It appears the authors were unaware of prior work, but unfortunately this submission cannot be published without major revision, because both the proposed problem (using learning algorithms to compute weightings to improve the predictions of the multi-model ensemble of AOGCMs) and the type of machine learning approach ("sequential learning algorithms") were already presented in previous work.

- C. Monteleoni, G. Schmidt, and S. Saroha. Tracking Climate Models. In NASA Conference on Intelligent Data Understanding (CIDU), pages 1–15, 2010.

- C. Monteleoni, G. Schmidt, S. Saroha, and E. Asplund. Tracking Climate Models. In Journal of Statistical Analysis and Data Mining: Special Issue: Best of CIDU 2010. Volume 4, Issue 4, pages 72–392, August 2011.
- S. McQuade and C. Monteleoni. MRF-Based Spatial Expert Tracking of the Multi-Model Ensemble. In New Approaches for Pattern Recognition and Change Detection, session at American Geophysical Union (AGU) Fall Meeting, 2013.

The first 2 are the most similar to the present submission; the 3rd is also highly related and was presented at AGU.

Moreover, while the stated task is identical to the first two works listed above, i.e. to use "sequential learning algorithms" (also known as "online learning with expert advice"), to combine the temperature projections of a CMIP ensemble (CMIP3 was used in the 2010 paper), the machine learning methods from the 2010 paper are more recent than the ones in the present submission and essentially subsume the methods used here. To see this, note that while the authors cite a textbook to refer to the algorithms (Cesa-Bianchi and Lugosi 2006, which we will refer to as CB06), the actual algorithms used are from the 1990's and earlier.
The algorithm denoted as EWA in the submission is originally due to:

- N. Littlestone and M. K. Warmuth. The weighted majority algorithm. In IEEE Symposium on Foundations of Computer Science, pages 256–261, 1989.
- V. Vovk. Aggregating strategies. In Proc. 3rd Annu. Workshop on Comput. Learning Theory, pages 371–383. Morgan Kaufmann, 1990.

The algorithm denoted as EGA in the submission is originally due to:

- Jyrki Kivinen and Manfred K. Warmuth: Exponentiated gradient versus gradient descent for linear predictors. Information and Computation 132(1):1-63, 1997.

In contrast, the algorithm in the 2010 paper (which we will hereby refer to as M10) is from the early 2000's:

- C. Monteleoni and T. Jaakkola. Online Learning of Non-stationary Sequences. In Advances in Neural Information Processing Systems 16. pages 1093–1100, 2003.

In the intervening time, the field of learning algorithms had advanced significantly (and subsequently it has advanced even further). In particular, relevant to the study of climate change and the problem in question, online learning algorithms were designed expressly for *non-stationary* data. The algorithms applied in the present submission were not designed to handle non-stationary data. Their performance guarantees (regret bounds) are with respect to the best fixed expert. However many climate scientists have shown that in any CMIP ensemble, some climate models perform better at some times and others at other times. An influential paper on online learning algorithms (extending algorithms with exponentiated weight updates for a variety of prediction loss functions) provided modified weight updates that allow for “switching” among "best" experts:

- M. Herbster and M. K. Warmuth, Tracking the best expert, Mach Learn 32 (1998), 151 – 178.

The algorithm used in M10 was a further extension from that work which actually learns the switching rate online (sequentially), simultaneous to updating the weights over the experts. It can still track the best fixed-expert as a special case, and in that sense it is significantly more flexible than the algorithms used in the present submission.

More recently, there have been several more machine learning approaches to the multi-model ensemble prediction problem proposed in the present submission. Note that the present submission treats geospatial locations separately. This was also the case in

M10, but subsequently, they did an extension to explicitly model geospatial neighborhood influence in the online learning weight updates. This is another highly related work to the present submission.

- S. McQuade and C. Monteleoni. Global Climate Model Tracking using Geospatial Neighborhoods. In Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence, Special Track on Computational Sustainability and AI, pp. 335–341, 2012.

The above was on a regular grid over latitude and longitude. However there has also been some recent work on learning the geospatial structure (need not be a grid) from the data such that regional features such as mountain ranges may cause some geospatially neighboring regions to have lesser spatial influence on one another.

- A. R. Goncalves, P. Das, S. Chatterjee, V. Sivakumar, and A. Banerjee. Global climate models combination using Multitask Sparse Structure Learning. Fourth Workshop on Understanding Climate Change Through Data, At National Center for Atmospheric Research (NCAR), Boulder, Colorado, USA, 2014.

The above is not a “sequential learning” technique. Additionally, there have been other kinds of machine learning algorithms applied to the problem, e.g.

- M. Ghafarianzadeh and C. Monteleoni. Climate Prediction via Matrix Completion. In Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence, Late-Breaking Papers Track, 2013.

Note, when the authors discuss using methods that predict “according to the best model,” they should see the related literature on learning algorithms. (In CB06, see Follow the Leader and Follow the Perturbed Leader).

In summary, the main problem with the present submission is its very strong similarity to previous work, along with a significant gap in the discussion of related work on using learning algorithms in climate predictions, in particular for the problem in question.

Technical corrections

There's a mistake in the equations. While it is more standard to state the weights in terms of the losses, L, than the regret, R, the authors chose to use R, and in the process a mistake was introduced into the equation for weights. When using L, there's a negative sign before the learning rate η in the exponent. However since R is defined as the loss of the expert *subtracted* from the loss of the learning algorithm, when using R in the exponent, there should *not* be negative sign. Intuitively, experts' weights should be *decreased* not increased, proportionally to their prediction losses. So the authors should either switch to L or else delete the negative in the exponent of Equation (3). (See pages 14, 72, and 304 of CB06).

Differences from past work, and positive aspects of the submission

On the positive side, the paper is very thoughtful and well-written. Moreover, there are also some novel experiments and ways of looking at the results, along with interesting insights from climate science. These should be emphasized in a revision.

- The input ensemble is from CMIP5 and has other differences from the ensembles used in several of the past works.
- Learning is done at a much higher spatial granularity than in M10 and some of the past works.
- The visualizations are different than M10 (although not all are different than all the past work on the problem, cf. Goncalves).
- There are some interesting findings, e.g. they point out that their results show that there is more model disagreement at polar regions.

- The authors use a variety of different metrics to evaluate their results instead of just prediction loss, e.g. RMSE (which however is closely related to the chosen prediction loss, also used in M10), as well as global, area-weighted uncertainty, and some interesting significance tests.
- The algorithm used in M10 has a per-time-iteration complexity that's on the order of a factor of $\sqrt{T}$ times that of the algorithm used in the present submission, where T is the desired number of temporal epochs (e.g. number of months for the whole experiment). This is for the optimized version of the algorithm in M10; one can also use lower complexity for slightly degraded results. This increased complexity is used in handling non-stationary data.

This is a difference, although not necessarily an advantage, from M10: The evaluations are done using a training and a test period. However this might not be the best technique, given that the observations and climate models exhibit non-stationarity. In such a setting, online learning algorithms are typically evaluated using *progressive validation error* and/or *cumulative prediction loss* (see M10).

Recommendations and suggested resources

It is very encouraging to see climate scientists using machine learning in their research. However this reviewer recommends resubmission, taking these major concerns into account. Given that there is already significant prior work in this area, the authors are strongly advised to re-think the framing of their submission, discussing past work, emphasizing the differences, and focusing much more on what is novel in their submission. Perhaps given the strong overlaps with past work, the authors will decide to run additional experiments with other learning algorithms with different properties, and compare them.

There is a field of interdisciplinary research using learning algorithms in climate science called "climate informatics" (see http://climateinformatics.org for a variety of links and a recent video tutorial surveying work in the field). There is an annual international climate informatics workshop (since 2011; the next workshop will be held be at NCAR in Colorado, in September 2015). These authors are doing work in this area and might benefit from this growing research community.

Interactive comment on Atmos. Chem. Phys. Discuss., 15, 7707, 2015.