Geosci. Model Dev. Discuss., 8, C1159–C1160, 2015
www.geosci-model-dev-discuss.net/8/C1159/2015/

Geoscientific Model Development Discussions

Interactive comment on "A global empirical system for probabilistic seasonal climate prediction" *by* J. M. Eden et al.

Anonymous Referee #1

Received and published: 17 June 2015

The title describes exactly the content of the paper: describing a statistical model and evaluating the seasonal probabilistic scores over the world. When I read this title, my first mind was "Yet another comparison of seasonal hindcast scores between numerical models and statistical methods." For reasons I will not develop here, I consider this type of contest as fair as comparing a car race with a horse race. But the content and the philosophy of the manuscript are completely different. The authors reduce strongly the score overestimation by selecting individually predictors amongst a few well known indices (not trying multiple combinations). They also use a progressive learning approach, closer to what would be a true forecast. They present their product as a complement of numerical operational forecasts, not as a challenger. This convinced me to propose this manuscript as suitable for publication, with a few minor corrections:

p2 lines 2 and 3: relies ... reliable

p3 line 10: of course, the model inadequacy wrt the true world induces systematic errors, but the main problem is elsewhere; if a model had just a cold bias, but would successfully predict the sequence of cold and warm seasons (which is measured by time correlation), one would be satisfied of it; conversely if a model has been carefully tuned and has a bias close to zero, but a very weak time correlation wrt observed seasons, one is not satisfied at all. So the critical point is the fact that, because a model has the wrong equations (assuming that the real world follows a small set of equations), its predictability is low. It is possible that, in addition, this model has biases, but the link bias-predictability is not very tight in practice.

p4 line 20: the non transferrability of empirical relations is a major criticism in the case of big climate change (e.g. 2100 RCP8.5). In the case of seasonal hindcasts spanning over the last 30 years, the stationarity hypothesis is acceptable. The major criticism is that these empirical relations could be partly based on coincidences of big events in the past which might not repeat in the future (I mean the coincidences, not the big events). There is thus a dilemma: longer time series, better robustness, lesser stationarity

p 10 line 16: why random sampling ? The reference forecast for calculating a skill score should be the forecast which minimizes this score in the absence of information: -average of the available past observations for RMSESS -distribution of the available past observations for CRPSSS

Interactive comment on Geosci. Model Dev. Discuss., 8, 3941, 2015.