DISCUSSION

Discussion of "A comparison of local and aggregated climate model outputs with observed data"*

A black eye for the *Hydrological Sciences Journal*

David Huard

Ouranos Inc., Montreal, Quebec H3A 1B9, Canada
huard.david@ouranos.ca

Recevied 24 February 2011; accepted 18 April 2011

Citation Huard, D. (2011) A black eye for the *Hydrological Sciences Journal.* Discussion of "A comparison of local and aggregated climate model outputs with observed data" by G.G. Anagnostopoulos *et al.* (2010, *Hydrol. Sci. J.* 55 (7), 1094–1110). *Hydrol. Sci. J.* **56**(7), 1330–1333.

Abstract A paper published by Anagnostopoulos *et al*. in volume 55 of the *Hydrological Sciences Journal* (*HSJ*) concludes that climate models are poor based on temporal correlation between observations and individual simulations. This interpretation hinges on a common misconception, that climate models predict natural climate variability. This discussion underlines fundamental differences between hydrological and climatological models, and hopes to clear misunderstandings regarding the proper use of climate simulations.

In a follow-up paper to Koutsoyiannis *et al*. (2008), Anagnostopoulos *et al.* (2010; Anagnostopoulos, Koutsoyiannis, Christofides, Efstratiadis and Mamassis, hereafter referred to as AKCEM), claim that climate models are poor on the basis of the low temporal correlation found between observed precipitation and temperature annual time series and model simulations during the 20th century. The idea of evaluating climate simulations initialized in the 19th century based on temporal correlation with observations at the yearly time scale is incongruous for anyone familiar with climate simulations and I could not understand how the paper had been accepted for publication. This discussion points out wrong assumptions used by AKCEM, and comments on the editorial review process, with the hope that it can somehow help correct a common misunderstanding about climate simulations.

The main issue with the AKCEM paper is that it is based on a false premise, namely that the selected climate simulations predict (forecast) climate in a deterministic sense. Climate models may indeed be used in a weather forecasting mode, and this is one way of evaluating their sub-grid scale parameterization at the sub-daily time scale (see for instance section 8.4.11 in Randall *et al*. 2007). Some are even used to "forecast" climate at the decadal scale (Keenlyside *et al*. 2008, Smith *et al*. 2007) using observed oceanic and atmospheric initial conditions, the oceanic inertia constraining the atmospheric model. However, these experiments are still considered highly experimental (Keenlyside and Ba 2010) and never claim to correlate with the inter-annual variability. Climate simulations included in IPCC's TAR and AR4 also make no pretence of predicting/forecasting weather or climate. As Smith *et al*. (2007) put it:

> Previous climate model projections of climate change accounted for external forcing from

*Anagnostopoulos, G.G., Koutsoyiannis, D., Christofides, A., Efstratiadis, A. and Mamassis, N. (2010) A comparison of local and aggregated climate model outputs with observed data. *Hydrological Sciences Journal*, 55 (7), 1094–1110.

ISSN 0262-6667 print/ISSN 2150-3435 online

http://dx.doi.org/10.1080/02626667.2011.610758
http://www.tandfonline.com

natural and anthropogenic sources but did not attempt to predict internally generated natural variability.

Evaluating climate models based on temporal correlations with observations is meaningful only if those models claim to forecast the year-to-year climate variations due to natural variability. The climate simulations analysed by AKCEM make no such claim, and the paper's main conclusion, that models are poor, is irrelevant.

DIFFERENCES BETWEEN HYDROLOGICAL AND CLIMATOLOGICAL MODELLING

A hydrological model is both deterministic and externally driven: its behaviour is completely dictated by the input time series: precipitation, temperature, etc. Although there are strong nonlinearities in hydrological processes, hydrological models are not chaotic. Consequently, given the same input data, hydrological simulations started with different initial conditions (e.g. water storage) all converge to approximately the same internal state and runoff after an appropriate spin-up period. For hydrological models, the output variability is entirely controlled by the sequence of input data, the external forcing.

Global climate models (GCM) are externally driven by solar radiation and planetary orbital parameters. An additional external forcing is the manmade emission of greenhouse gases (GHG) and aerosols. None of these external forcings, however, can explain the inter-annual variability present in all climate variables. Variability at the annual and decadal scales emerges spontaneously from the dynamics of the climate system and is only weakly influenced by external forcing (massive volcanic eruptions are an exception). At the multi-decadal scale, variability is caused by a mix of natural variability and changes in external forcing conditions. Murphy *et al*. (2009) provide a clear and crisp discussion around these concepts.

The natural variability of the climate system is largely chaotic (Lorenz 1963). This fundamental unpredictability, combined with the structural instability of climate models (McWilliams 2007), means that two climate simulations started with only slightly different initial conditions, or slightly different parameterizations, will eventually diverge: their time correlation goes to zero. So if two climate simulations run under similar conditions are not expected to show a correlation after a decade or two, it is absurd to expect inter-annual correlation between observations and individual climate simulations started in 1860! Nevertheless, AKCEM interpret this lack of correlation as evidence that models are poor. In reality, it is evidence that there is chaotic natural variability at all time scales (Scherrer 2010).

EXTRACTING THE EXTERNAL FORCING SIGNAL

One of other metrics used by AKCEM to evaluate model performance is the correlation between the 30-year running mean of simulations and observations. This case is much more interesting, since we expect external forcing by GHGs to play a role at that time scale, and, thus, to explain a portion of the observed variability. Note that this is very different from saying that climate models predict climate; *under constant external forcing*, TAR and AR4 simulations have no predictive skill whatsoever on the chronology of events beyond the annual cycle. A climate *projection* is thus not a prediction of climate, it is an experiment probing the model's response to change in GHG concentrations. In fact, early climate change experiments only looked at differences between steady states under present conditions and a doubling of CO_2 (Manabe and Wetherald 1975). Current transient experiments probe how fast and exactly how the climate reaches a new steady state, but the inter-annual variability is still mostly the result of unpredictable natural variability.

To evaluate the skill of climate models at capturing the effect of changes in GHG concentrations on climate, we need to filter out the natural variability. To do this, we typically average multiple members of an ensemble of simulations. The hypothesis is that the climate components that are due to natural variability will average out and leave only the response of models to the external forcing (see Wang *et al*. 2007 for an example of such model evaluation). AKCEM apparently did not realize this and averaged correlation coefficients computed with individual simulations, instead of computing the correlation from averaged simulations.

Another comparatively minor problem is that the number of independent 30-year samples in the time series used by AKCEM is rather small to compute a meaningful correlation coefficient: assuming a series has 120 years and using Fisher's r-to-z transformation, the 95% confidence interval for a Pearson coefficient of 0.7 computed with four samples is [–0.8;1]. Instead, AKCEM applied a moving 30-year averaging window to the annual time series. While this approach provides more numbers to work with (N – 30), the variability of the resulting time series is not exclusively due to the 30-year variability, but also includes annual variability, which again is largely chaotic.

AKCEM expected individual models to show some skill in predicting multi-decadal climate variations. They do, but their skill is limited to the small fraction of climate's variability driven by external forcing. To evaluate model performance, it is fundamental to extract the model's response to the external forcing from the background natural variability (Randall *et al*. 2007). Failing to do this, AKCEM have merely shown that climate models display chaotic behaviour at small and long time scales, not that they are poor.

THE DECISION TO PUBLISH

So why was the paper published even though its methodology is naive and the conclusion misleading? I asked Frances Watkins from the *HSJ* Editorial Office for a copy of the anonymous reviews, the author's response and the editor's decision letter. The authors, editors and reviewers all agreed to make these documents available and with those I could make some sense of what happened.

The paper received three evaluations. Reviewer A provided a solid review which identified unsubstantiated or false claims and methodological shortcomings. The evaluation included a comment on the general lack of rigour of the paper with a recommendation not to publish. Reviewer B rated the paper as "Very good to excellent" and made three superficial suggestions for improvements. Reviewer C rated the paper as "Poor to fair" and specifically stated: "*This paper is misleading as it is based on a wrong assumption related to the climate system predictability*." Reviewer C also criticized the methodology as inappropriate and recommended the paper be rejected outright.

Faced with these reviews, Dr Kundzewicz in his decision letter writes he heeded advice from the late Stephen Schneider (editor of Climatic Change): "*When I get a paper that generates controversy and splits reviewer advice, I look to be sure that it is mostly differing philosophy rather than technical errors that underlie the dispute*."

While I certainly agree with this guiding principle, reviewers A and C rejected the paper on technical and methodological grounds, not philosophy. Their review highlights a general lack of rigour, methodological issues and factual errors obvious to anyone familiar with climate science. Although there may be philosophical differences between AKCEM and the critics, this should not be an excuse to dismiss methodological issues. In my experience as author and reviewer, strong opposition to publication by two reviewers out of three generally leads either to rejection, or, if the editor feels the paper has merit, to an additional review. In this case, however, it seems that the critics were not taken seriously, or even understood. Indeed, the editorial piece (Kundzewicz and Stakhiv 2010) indicates the editor shares the same misguided assumptions about climate simulations as AKCEM and there is little hope that an additional review would have made any difference. This is in my view a black eye for *HSJ* coming out as lacking the discrimination required to identify poor science.

As someone who went from quantum mechanics to hydrology to sea ice modelling and finally climatology, I can say with confidence that experience in a field does not translate as is to other disciplines. The same words do not carry the same meaning and the basic assumptions and expectations about models differ quite a lot. Collaboration with other scientific communities requires attention to detail, care and introspection, but more importantly, mutual respect and trust in the professionalism of our peers. I would like to praise Frances Watkins, Dr Kundzewicz, Dr Koutsoyiannis, Dr Wilby and the anonymous reviewers for their openness and willingness to discuss the paper. I would also like to thank my colleagues for their suggestions and support.

REFERENCES

Anagnostopoulos, G.G., Koutsoyiannis, D., Christofides, A., Efstratiadis, A. and Mamassis, N. (2010) A comparison of local and aggregated climate model outputs with observed data. *Hydrological Sciences Journal*, 55 (7), 1094–1110. doi:10.1080/02626667.2010.513518.

Keenlyside, N.S. and Ba, J., 2010. Prospects for decadal climate prediction. *WIREs Climate Change*, 1 (5), 627–635. ISSN 1757-7799. doi:10.1002/wcc.69. http://dx.doi.org/10.1002/wcc.69.

Keenlyside, N.S., Latif, M., Jungclaus, J., Kornblueh, L. and Roeckner, E., 2008. Advancing decadal-scale climate prediction in the North Atlantic sector. *Nature*, 453 (7191), 84–88. doi:10.1038/nature06921.

Koutsoyiannis, D., Efstratiadis, A., Mamassis, N. and Christofides A., 2008. On the credibility of climate predictions. *Hydrological Sciences Journal*, 53 (4), 671–684.

Kundzewicz, Z.W. and Stakhiv, E.Z., 2011. Are climate models "ready for prime time" in water resources management applications, or is more research needed? *Hydrological Sciences Journal*, 55 (7), 1085–1089. doi:10.1080/02626667.2010.513211.

Lorenz, E.N., 1963. Deterministic nonperiodic flow. *Journal of Atmospheric Sciences*, 20, 130–141.

Manabe, S. and Wetherald, R.T., 1975. The effects of doubling the CO_2 concentration on the climate of a general circulation model. *Progress in Physical Geography*, 25, 385–387.

McWilliams, J.C., 2007. Irreducible imprecision in atmospheric and oceanic simulations. *Proceedings of the National Academy of Sciences*, 104 (21), 8709–8713. doi:10.1073/pnas.0702971104.

Murphy, J., *et al*., 2009. *UK Climate projections science report: Climate change projections*. Exeter, UK: Met Office Hadley Centre.

Randall, D.A., *et al.,* 2007. Climate models and their evaluation. Chapter *in: 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 : Cambridge University Press. Available at: http://ipcc-wg1.ucar.edu/wg1/wg1-report.html.

Scherrer, S.C., 2010. Present-day interannual variability of surface climate in CMIP3 models and its relation to future warming. *International Journal of Climatology*, doi:10.1002/joc.2170.

Smith, D.M., *et al.,* 2007. Improved surface temperature prediction for the coming decade from a global climate model. *Science*, 317 (5839), 796–799. doi:10.1126/science.1139540.

Wang, M., *et al*., 2007. Intrinsic *versus* forced variation in coupled climate model simulations over the Arctic during the twentieth century. *Journal of Climate*, 20 (6), 1093–1107. doi:10.1175/JCLI4043.1.