Hydrological Sciences Journal
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t911751996

Evaluating climate model outputs for hydrological applications

R. L. Wilby[a]
[a] Centre for Hydrological and Ecosystem Science, Loughborough University, Loughboroough, UK

Online publication date: 13 October 2010

To cite this Article Wilby, R. L.(2010) 'Evaluating climate model outputs for hydrological applications', Hydrological Sciences Journal, 55: 7, 1090 — 1093

To link to this Article: DOI: 10.1080/02626667.2010.513212

URL: http://dx.doi.org/10.1080/02626667.2010.513212

OPINION

Evaluating climate model outputs for hydrological applications

R. L. Wilby

Centre for Hydrological and Ecosystem Science, Loughborough University, Loughboroough, LE11 3TU, UK
r.l.wilby@lboro.ac.uk

Open for discussion until 1 April 2011

Citation Wilby, R. L. (2010) Evaluating climate model outputs for hydrological applications – Opinion. *Hydrol. Sci. J.* **55**(7), 1090–1093.

How useful are climate models for long-term water management and adaptation planning? Or put another way: as hydrologists, what features would we like to see represented well in regional climate simulations? These questions are exercising both scientific and policy-making communities as the prospect of unavoidable climate change grows with each failed attempt to set international targets for greenhouse gas emissions. They also expose a more profound conceptual divide between those advocating a *scenario-led* approach to adaptation, and those who have concluded that scenarios are better used for sensitivity testing and adaptation *options appraisal* (Wilby & Dessai, 2010). This note addresses both applications of climate models: first, by offering criteria that may help evaluate climate model skill from a water management perspective; second, by closing with remarks about the utility of scenario-led approaches for adaptation in practice.

Hydrological processes operate across at least nine orders of magnitude in both space and time (Blöschl & Sivapalan, 1995). Attendant management scales are equally diverse: spanning minutes for real-time flood protection through to multiple decades for water provision. Therefore, it is unlikely that any single review of climate model outputs will be sufficiently broad to cover all scales of hydrological interest. We should also keep in mind that the original intent of coupled ocean/atmosphere general circulation models (OA/GCMs) was to assess the global consequences of different emission pathways and, only relatively recently, how we might adapt at regional and local scales (Schiermeier, 2007). Downscaling techniques are one manifestation of wider efforts to bridge these discordant scales under the general mantra of *climate risk information at the scale required for decision making* (see for example, Fowler *et al.*, 2007).

With the above points in mind, five principles are recommended for those selecting climate models for hydrological applications:

1. *Quantify the uncertainty in the observed data used for model evaluation*. In any comparison study there should always be healthy scepticism about the quality and homogeneity of the observed data used to gauge model performance. It is well known that incremental or step changes in meteorological records can arise from changes in site, instrumentation, observing or recording practices, site characteristics, or sampling regime (Kundzewicz & Robson, 2004). Observing networks may be spatially biased towards lower-elevation, urban or coastal locations. Discharge records may be influenced by non-climatic factors including changes in land cover and management, urbanization, river regulation, water abstraction and effluent returns, or flood-flows by-passing gauging structures. Trends found in shorter series may cease to be significant when the influence of outliers (at the start or the end of the record), or multi-decadal variability have been taken into account. Ideally, confidence intervals should be provided for all observations to reflect the time- and space-varying density, completeness and homogeneity of data used for benchmarking models.
2. *Compare like with like*. Climate model output is most readily available at monthly time scales and at spatial resolutions of 25 to 300 km for selected

ISSN 0262-6667 print/ISSN 2150-3435 online

doi: 10.1080/02626667.2010.513212
http://www.informaworld.com

levels in the atmosphere. Where daily quantities are employed, beware that some earlier GCMs have 360 days in a notional "year". This means that we must be realistic about our expectations of the model and use observations that have been aggregated to comparable scales (e.g. New *et al.*, 2002). Clearly, the GCM is incapable of replicating sub-grid-scale processes involving clouds, complex topography, land-water or surface cover boundary effects. For example, the smoothed orography of GCMs cannot produce known variations in lapse rates, or local snow–ice feedbacks for sites near the 0°C isotherm (Pepin & Lundquist, 2008). Therefore, it makes little sense to compare gridded climate model output directly with point measurements of rainfall and temperature without some form of area-reduction (e.g. Fowler *et al.*, 2005), scaling (e.g. Osborn & Hulme, 1997), or bias-correction (e.g. Schmidli *et al.*, 2006). There should also be an upper limit to the domain size used for aggregation. For instance, 20th century seasonal and annual precipitation totals for North America as a whole would merge quite divergent trends over the southwest (drying) and northeast (wetter) (Trenberth *et al.*, 2007). In this case, correlation patterns between observed and modelled precipitation change would help to identify regions with greatest and least skill.

3. *Select indicators of performance relevant to intended hydrological applications*. The most pressing climate information needs in water and agriculture sectors are for high-frequency weather (extremes) and lower-frequency climate (variability) over years to decades (WMO, 2010). As noted above, GCMs will be constrained in their ability to represent realistic weather phenomena until global simulations can be performed sub-hourly, at spatial resolution of a few kilometres. In the meantime, it is more sensible to compare observations with downscaled weather extremes than raw GCM output (e.g. Haylock *et al.*, 2006). However, there is scope to evaluate model skill at simulating observed multi-scale variability using metrics that are relevant to hydrological assessments (Barbosa *et al.*, 2009; Johnson & Sharma, 2009). For example, climate model experiments with observed natural forcing can reproduce the abrupt decline in precipitation across the Sahel since the 1970s (Held *et al.*, 2005), but have been less skillful at mimicking known teleconnections between Pacific sea-surface temperature anomalies and the South Asian monsoon (Annamalai *et al.*, 2007). Above all, it is important to assess model skill using multiple diagnostics because inter-comparison projects show that no single model consistently outperforms all others (Gleckler *et al.*, 2008). This further underlines the need to evaluate large ensembles of GCMs as there is always a danger that a small sample of models will not be representative. Greater insight is also gained when the physical basis for any discrepancies is analysed, such as the recognized over-sensitivity of modelled precipitation to variation in humidity (Wilby & Wigley, 2000).
4. *Evaluate climate models relative to other components of hydrological uncertainty.* Climate model outputs are known to vary with the conditions used to initialize each experiment, with different parameterizations for sub-grid processes, and due to internally generated multi-decadal variability. Increasingly, these factors are being analysed via large-ensemble GCM-modelling efforts such as *ClimatePrediction.net* (Stainforth *et al.*, 2005) and UKCP09 (Murphy *et al.*, 2009), thereby offering probabilistic climate projections for regional water planning. Although understanding of downscaling uncertainty has improved thanks to projects such as ENSEMBLES, PRUDENCE and STARDEX, relatively little is known about the significance of impact model uncertainty. This means that the overall uncertainty is underestimated and biased towards climatic elements. Calls for the research community to populate so-called "hyper-matrices" with results from different permutations of forcing scenario, climate model, initial conditions, downscaling and region appear to overlook impact model uncertainty (Giorgi *et al.*, 2008). However, for the next few decades hydrological model uncertainty will be a larger source of overall uncertainty than the emissions pathway (Wilby & Harris, 2006). Therefore, if we are serious about characterizing the range of uncertainty in future hydrological projections, we need to look beyond evaluating climate models alone.
5. *Test combined climate, downscaling and hydrological model skill using near-term applications.* Many model evaluations are predicated on the grounds that evidence of skill for the present climate translates into higher confidence about future climate and associated impacts – but this is by no means guaranteed (Knutti, 2008). As noted before, such experiments are more informative when some

physical insight into model behaviour is gained. Returning to the point about choosing impact-relevant diagnostics of model performance, low skill in simulating high-intensity precipitation events need not preclude the use of the same model if it produces realistic sequences of multi-seasonal drought. Real-time and seasonal forecasting with coupled atmosphere–catchment models offer useful analogues of climate change scenarios with the advantage of testing under near-term conditions. For example, Wood *et al.* (2002) used a quantile-quantile method to downscale coarse-resolution seasonal forecasts into basin-scale information for operational river flow outlooks, whilst Leung & Qian (2009) examined the influence of atmospheric rivers and land surface conditions on heavy precipitation and flooding in the western USA Hence, even when climate model and derived outputs compare unfavourably with observations, useful information might still be gained about the *causes* of the discrepancy, thereby stimulating further research and development.

Even if we could build perfect climate models, uncertainty about future economic and demographic pathways, natural forcings by solar and volcanic activity, and a host of non-climatic pressures, mean that regional hydrological projections would still be highly uncertain. In other words, *characterizing* uncertainty through concerted scientific action may be a tractable proposition, but there appears to be no immediate prospect of *reducing* uncertainty in the risk information supplied to decision makers. There is now widespread acceptance that the assumptions of stationary conditions that have underpinned past flood, water and conservation management are no longer helpful (Milly *et al.*, 2008). However, the tendency to adopt a "predict and provide" strategy still prevails in the shape of scenario-led adaptation planning. It remains to be seen whether theoretical advances in probabilistic projections and downscaling really can shape practical adaptation responses.

Such views are leading some to conclude that "top down" frameworks might not facilitate the most effective use of climate model outputs for adaptation planning (Dessai *et al.*, 2005). A few studies have experimented with sensitivity testing and adaptation options appraisal for flood risk, water quality and resource management. In some cases, climate risk information is used to bracket the range of changes for sensitivity testing. For example, Prudhomme *et al*. (2010) developed a method for testing the effectiveness of UK Government safety margins for flood protection under a wide range of climate change scenarios. Others have investigated future water security and water quality by benchmarking different adaptation options against business-as-usual outcomes (e.g. Whitehead *et al*., 2006; Lopez *et al*., 2009). The novel feature of these studies is that they move away from thinking that there is a single optimum solution towards designing adaptation pathways that are robust despite the uncertain outlook. This does not obviate the need for reliable climate models, but it does shift the emphasis onto considering first what is socially, economically and technically feasible. Paradoxically, the better we characterize the breadth of uncertainty in future regional climates, the less we are likely to depend on this information in a deterministic sense.

REFERENCES

Annamalai, H., Hamilton, K. & Sperber, K. R. (2007) The South Asian summer monsoon and its relationship with ENSO in the IPCC AR4 simulations. *J. Climate* **20**, 1071–1092.

Barbosa, S., Silva, M. & Fernandes, M. (2009) Multi-scale variability patterns in NCEP/NCAR reanalysis sea-level pressure. *Theoret. Appl. Climatol.* **96**, 319–326.

Blöschl, G. & Sivapalan, M. (1995) Scale issues in hydrological modelling: a review. *Hydrol Processes.* **9**, 251–290.

Dessai, S., Lu, X. & Risbey, J. S. (2005) On the role of climate scenarios for adaptation planning. *Global Environ. Change* **15**, 87–97.

Fowler, H. J., Blenkinsop, S. & Tebaldi, C. (2007) Linking climate change modelling to impact studies: recent advances in downscaling techniques for hydrological modelling. *Int. J. Climatol.* **27**, 1547–1578.

Fowler, H. J., Ekström, M., Kilsby, C. G. & Jones, P. D. (2005) New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 1. Assessment of control climate. *J. Hydrol.* **300**, 212–233.

Giorgi, F., Diffenbaugh, N. S., Gao, X. J., Coppola, E., Dash, S. K., Frumento, O., Rauscher, S. A., Remedio, A., Seidou Sanda, I., Steiner, A., Sylla, B. & Zakey, A. S. (2008) The regional climate change hyper-matrix framework. *Eos* **89**, 445–446.

Gleckler, P. J., Taylor, K. E. & Doutriaux, C. (2008) Performance metrics for climate models. *J. Geophys. Res.* **113**, D06104.

Haylock, M. R., Cawley, G. C., Harpham, C., Wilby, R. L. & Goodess, C. M. (2006) Downscaling heavy precipitation over the UK: a comparison of dynamical and statistical methods and their future scenarios. *Int. J. Climatol.* **26**, 1397–1415.

Held, I. M., Delworth, T. L., Lu, J., Findell, K. L. & Knutson, T. R. (2005) Simulation of Sahel drought in the 20th and 21st centuries. *Proc. Nat. Acad. Sci. USA* **102**, 17891–17896.

Johnson, F. & Sharma, A. (2009) Measurement of GCM skill in predicting variables relevant for hydroclimatological assessments. *J. Climate* **22**, 4373–4382.

Knutti, R. (2008) Should we believe model predictions of future climate change? *Phil. Trans. Roy. Soc. A* **366**, 4647–4664.

Kundzewicz, Z. W. & Robson, A. J. (2004) Change detection in hydrological records – a review of the methodology. *Hydrol. Sci. J.* **49**(1), 7–19.

Leung, L. R. & Qian, Y. (2009) Atmospheric rivers induced heavy precipitation and flooding in the western U.S. simulated by the WRF regional climate model. *Geophys. Res. Lett.* **36**, L03820.

Lopez, A., Fung, F., New, M., Watts, G., Weston, A. & Wilby, R. L. (2009) From climate model ensembles to climate change impacts: a case study of water resource management in the South West of England. *Water Resour. Res.* **45**, W08419.

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P. & Stouffer, R. J. (2008) Stationarity is dead: whither water management? *Science* **319**, 573–574.

Murphy, J. M., Sexton, D. M. H., Jenkins, G. J., Booth, B. B. B., Brown, C. C., Clark, R. T., Collins, M., Harris, G. R., Kendon, E. J., Betts, R. A., Brown, S. J., Humphrey, K. A., McCarthy, M. P., McDonald, R. E., Stephens, A., Wallace, C., Warren, R., Wilby, R. & Wood, R. A. (2009) *UK Climate Projections Science Report: Climate Change Projections*. Exeter: Met Office Hadley Centre.

New, M. G., Lister, D., Hulme, M. & Makin, I. (2002) A high-resolution data set of surface climate for terrestrial land areas. *Climate Res.* **21**, 1–25.

Osborn, T. J. & Hulme, M. (1997) Development of a relationship between station and grid-box rainday frequencies for climate model evaluation. *J.Climate* **10**, 1885–1908.

Pepin, N. C. & Lundquist, J. D. (2008) Temperature trends at high elevations; patterns across the globe. *Geophys. Res. Lett.* **35**, L14701.

Prudhomme, C., Wilby, R. L., Crooks, S., Kay, A. L. & Reynard, N. S. (2010) Scenario-neutral approach to climate change impact studies: application to flood risk. *J. Hydrol.* **390**, 198–209.

Schiermeier, Q. (2007) Get practical, urge climatologists. *Nature* **448**, 234–235.

Schmidli, J., Frei, C. & Vidale, P. L. (2006) Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods. *Int. J. Climatol.* **26**, 679–689.

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. & Allen, M. R. (2005) Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature* **433**, 403–406.

Trenberth, K. E., Jones, P. D., Ambenje, P. Bojariu, R., Easterling, D., Klein Tank, A., Parker, D., Rahimzadeh, F., Renwick, J. A., Rusticucci, M., Soden, B. & Zhai, P. (2007) Observations: surface and atmospheric climate change. 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 (S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor & H. L. Miller, eds), Cambridge: Cambridge University Press.

WMO (World Meteorological Organization) (2010) White paper on climate applications/services theme for WCRP. Joint Scientific Committee, Antalya, Turkey. http://www.wmo.int/wcrpevent/jsc31/documents/jsc-31WPApplications_2.4.pdf

Whitehead, P. G., Wilby, R. L., Butterfield, D. & Wade, A. J. (2006) Impacts of climate change on nitrogen in a lowland chalk stream: an appraisal of adaptation strategies. *Sci. Total Environ.* **365**, 260–273.

Wilby, R. L. & Dessai, S. (2010) Robust adaptation to climate change. *Weather* **65**, 180–185.

Wilby, R. L. & Harris, I. (2006) A framework for assessing uncertainties in climate change impacts: low flow scenarios for the River Thames, UK. *Water Resour. Res.* **42**, W02419. doi:10.1029/2005WR004065.

Wilby, R. L. & Wigley, T. M. L. (2000) Precipitation predictors for downscaling: observed and general circulation model relationships. *Int. J. Climatol.* **20**, 641–661.

Wood, A. W., Maurer, E. P., Kumar, A. & Lettenmaier, D. P. (2002) Long range experimental hydrologic forecasting for the eastern US. *J. Geophys. Res.* **107**(D20), 4429.