Model independence and the representation of model space as a source of uncertainty

Gab Abramowitz[1,2] and Hoshin Gupta[3]

[1]*UNSW Climate Change Research Centre*
[2]*Centre for Australian Weather and Climate Research*
[3]*University of Arizona*

This talk explores the idea of a "model space" as the space of all possible models for a given prediction problem. While we may be familiar with the notion of uncertainty in model parameters, states or inputs, uncertainty associated with the model we use is clearly a more difficult problem. It's easy for us to conceptualise the parameter, state or input spaces of a model because they are real number spaces (Figure 1). Uncertainty in these spaces can be easily defined with probability density functions (PDFs) and, through either frequentist or Bayesian techniques, propagated through the model to the (real number) output space. A well known example of this is climateprediction.net (Stainforth et al, 2005), where thousands of perturbed parameter runs of HadCM3 were used to estimate the climate response to anthropogenic CO_2 forcing. While a pdf of model output behaviour obtained in this way might represent the probability of a particular model outcome, it is clearly not equivalent to the probability of the event actually occurring. This distinction is precisely because of model space uncertainty.

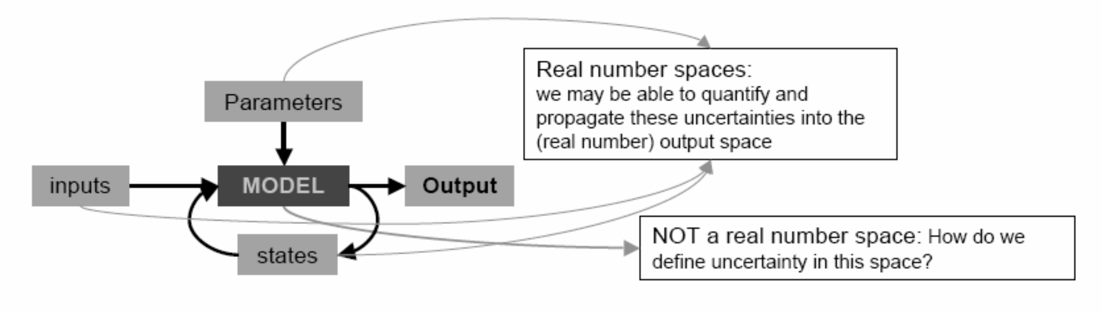

Figure 1: A typical systems representation of a model showing input, parameter, state and output spaces as real number spaces.

The most common approach to overcoming this issue is the use of multi-model ensemble simulations (e.g., Houghton et al, 2001; Gillet et al, 2002). The rationale is that different modelling groups produce quite different models and so provide independent estimates, but this independence is rarely if ever quantified (Tebaldi and Knutti, 2007). Models may be different for a variety of reasons:

- Different processes may be included. This can be thought of as a difference in *perceptual models* – the processes we see or imagine are part of the natural system.
- Different process interactions may be included: This can be thought of as a difference in *conceptual models* – the nature of the relationship between processes.
- Different symbolic representation of processes which are included – that is, differences in the translation from conceptual model to *mathematical model*.
- Differences in numerical or scale implementation, precision – the *numerical model*.

Even with these potential differences in mind, it's hard to judge how independent models need to be. The hope is that an ensemble samples the model space broadly enough so that in most measures of performance, the multi-model ensemble should be unbiased relative to observations. To date however, we have no technique to define dependence or independence of models, and so no way of knowing what 'spanning' the model space may mean.

We attempt to formalise the idea of the model space by proposing a distance measure, or metric, for this space. That is, a way of measuring the distance between two models. If we can successfully cast the model space as a metric space in the mathematical sense, this may allow us to treat the model space in a similar statistical framework to real number spaces such as parameter, input and state space. To begin with however, we simply use this distance measure as a proxy for model dependence.

To illustrate the proposed model space metric, we use land surface models (LSMs), specifically the prediction of net ecosystem exchange of CO_2 (NEE) as a function of site/grid cell meteorology. Measured values of meteorological data are taken from 13 flux tower sites in Europe and the USA from a variety of biomes, totalling about 40 site years, and used to drive the LSMs.

The first step is to divide all the time steps of meteorological data into discrete groups (henceforth 'nodes'), each containing similar conditions. A very simple way to do this, for example, might be air temperature > (or <) 15C and relative humidity > (or <) 50% - this would provide us with four climatologically divided nodes. For this work we chose to use a self-organising map (SOM, Kohonen, 1989) to divide meteorological data into 9 nodes. The mean values of four meteorological variables for each of the 9 nodes are shown in Figure 2.

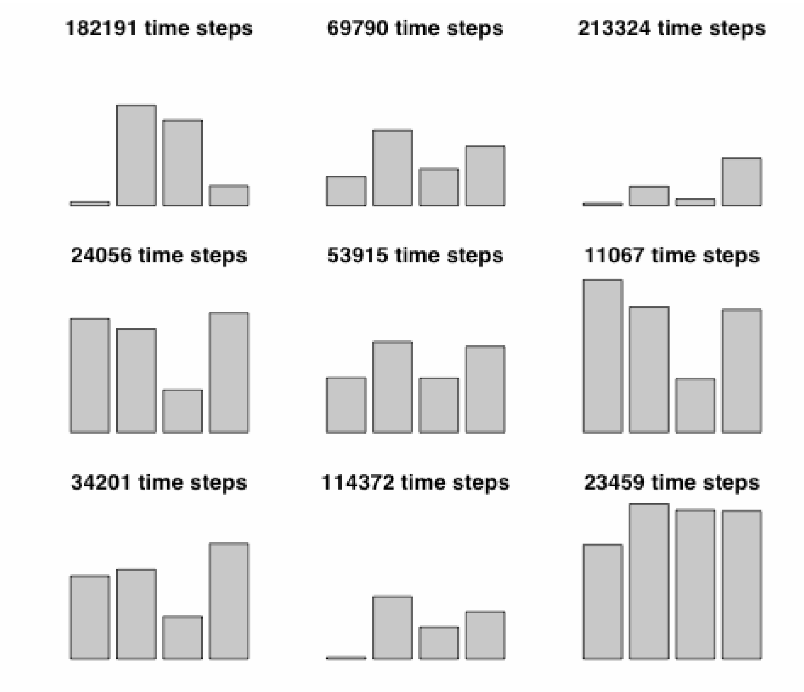

Figure 2: Average values for four LSM input variables: SW down, air temperature, specific humidity and windspeed for time steps belonging to each node of a 9-node Self Organising Map. This meteorological data, used to drive the LSMs, was sourced from 13 eddy covariance flux tower sites across two continents.

For each time step belonging to one of these nodes, we extract the corresponding NEE values from the time series of predictions by two LSMs: CABLE (Kowalczyk et al, 2006) and ORCHIDEE (Krinner et al, 2005). For each LSM, we construct a PDF of the NEE predictions associated with each node. These PDFs are shown for each node in Figure 3.

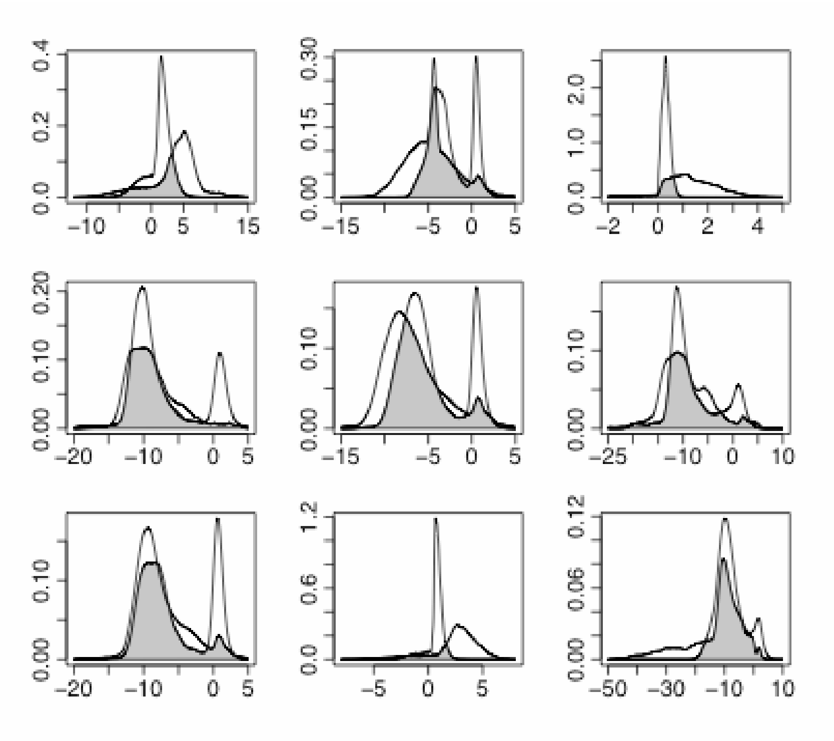

Figure 3: Probability density functions of CABLE's and ORCHIDEE's prediction of NEE for each SOM node (conditions and the number of data of each node are shown in previous figure). Note that abscissa units are µmol/m^2/s and that nodes have different scales to show relative model differences.

The grey region in each panel of Figure 3 represents the overlap of CABLE's and ORCHIDEE's PDF of NEE predictions for the particular set of meteorological conditions represented by a SOM node. If we represent this overlap region for the nth node as O_n and the number of time steps belonging to the nth node as a_n, then we define the distance between two models m_1 and m_2 as:

$$d(m_1,m_2)=1-\sum_{n=1}^{N}\left[\frac{a_n}{A}\bullet O_n\right]$$

where *A* is the total number of time steps (here 726384) and *N* is the total number of nodes (here 9). That is, we simply add all areas associated with the shaded regions, weighted by the number of time steps belonging to each node. Since these are PDFs, this sum is between 0 and 1. While we have not proved it to be the case, it seems plausible that this measure of the model space may satisfy the metric space axioms: *d* ³ 0 and it will only be zero when all density functions overlap perfectly, so *d*?m_1,m_2?=0 if and only if m_1=m_2. It should also be clear that *d*?m_1,m_2?=*d*?m_2,m_1?. No immediate counter examples to the triangle inequality *d*?m_1,m_2??*d*?m_2,m_3?=*d*?m_1,m_3? seem apparent either.

Using this metric, any binary decision about which models are dependent or independent is simply a matter of choosing a 'dependence radius' – models separated by less than this radius could be considered to be dependent, those separated by more to be independent. In most multi-model ensemble applications, however, this information would be used as a separate model property, alongside performance, in the construction of model weights. The process might proceed as follows:

1. Construct model weights using a (collection of) performance measures.
2. Reduce or increase a model's weight according to its distance to all other models. For example: w_1 = P_1.(*S* *d*(m_1,m_i))/(w_1+…+w_n) where P_1 is the performance of model 1.

The conceptualisation of the model space as a metric space presented here may give us a much needed tool that could allow us to assess the relationship between model spread and prediction uncertainty. While we are clearly at a very early stage of understanding model independence, it should be obvious that it needs to be a separate consideration to model performance when constructing an unbiased model ensemble.

Gillett, N. P., F. W. Zwiers, A. J. Weaver, G. C. Hegerl, M. R. Allen, and P. A. Stott, 2002. Detecting anthropogenic influence with a multi-model ensemble, *Geophysical Research Letters*, 29, 1970, doi:10.1029/2002GL015836.

Houghton, J. T., Y. Ding, M. Nogua, D. Griggs, P. Vander Linden, K. Maskell (eds.) 2001. Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, U.K.

Kohonen, T. 1989. Self-Organization and Associative Memory. Springer-Verlag.

Kowalczyk, E. A., Y. P. Wang, R. M. Law, H. L. Davies, J. L. McGregor and G. Abramowitz, 2006. The CSIRO Atmosphere Biosphere Land Exchange (CABLE) model for use in climate models and as an offline model CSIRO Marine and Atmospheric Research paper 013, www.cmar.csiro.au/e-print/open/kowalczykea_2006a.pdf 264

Krinner, G., N. Viovy, N. de Noblet-' XFRXGUH, J. 2J PH, J. Polcher, P. Friedlingstein, P. Ciais, S. Sitch and I. C. Prentice, 2005. A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*, 19(GB1015).

Stainforth, D. A., T. Aina, C. Christensen, M. Collins, N. Faull, D. J. Frame, J. A. Kettleborough, S. Knight, A. Martin, J. M. Murphy, C. Piani, D. Sexton, L. A. Smith, R. A. Spicer, A. J. Thorpe & M. R. Allen, 2005. Uncertainty in predictions of the climate response to rising levels of greenhouse gases, *Nature*, 433, pp.403-406.

Tebaldi, C. and R. Knutti, 2007. The use of the multimodel ensemble in probabilistic climate projections, *Philosophical Transactions of the Royal Society A*, 365, 2053-2075, doi:10.1098/rsta.2007.2076.