Geophysical Research Abstracts
Vol. 18, EGU2016-1731-1, 2016
EGU General Assembly 2016

Information-Theoretic Perspectives on Geophysical Models

Grey Nearing
NASA GSFC, Hydrologic Sciences Lab, Maryland, United States (grey.s.nearing@nasa.gov)

To test any hypothesis about any dynamic system, it is necessary to build a model that places that hypothesis into the context of everything else that we know about the system: initial and boundary conditions and interactions between various governing processes (Hempel and Oppenheim, 1948, Cartwright, 1983). No hypothesis can be tested in isolation, and no hypothesis can be tested without a model (for a geoscience-related discussion see Clark et al., 2011).

Science is (currently) fundamentally reductionist in the sense that we seek some small set of governing principles that can explain all phenomena in the universe, and such laws are ontological in the sense that they describe the object under investigation (Davies, 1990 gives several competing perspectives on this claim). However, since we cannot build perfect models of complex systems, any model that does not also contain an epistemological component (i.e. a statement, like a probability distribution, that refers directly to the quality of of the information from the model) is falsified immediately (in the sense of Popper, 2002) given only a small number of observations. Models necessarily contain both ontological and epistemological components, and what this means is that the purpose of any robust scientific method is to measure the amount and quality of information provided by models. I believe that any viable philosophy of science must be reducible to this statement.

The first step toward a unified theory of scientific models (and therefore a complete philosophy of science) is a quantitative language that applies to both ontological and epistemological questions. Information theory is one such language: Cox' (1946) theorem (see Van Horn, 2003) tells us that probability theory is the (only) calculus that is consistent with Classical Logic (Jaynes, 2003; chapter 1), and information theory is simply the integration of convex transforms of probability ratios (integration reduces density functions to scalar metrics) (Csiszár, 1972).

Fundamentally, models can only translate existing information – they cannot create information. That is, all of the information about any future (or otherwise unobserved event) is contained in the initial and boundary conditions of whatever model we will use to predict that phenomena (Gong et al., 2013). A model simply tells us how to process the available information in a way that is as close to isomorphic with how the system itself processes information. As such, models can only lose or corrupt information because at best a model can only perfectly extract all information contained in its input data; this is a theorem called the Data Processing Inequality (Cover and Thomas, 1991), and this perspective represents a purely ontological treatment of information in models. In practice, however, models provide information to scientists about how to translate information, and in this epistemic sense, models can provide positive quantities of information.

During engineering-type efforts, where our goal is fundamentally to make predictions, we would measure the (possibly positive) net epistemic information from some hypothesized model relative to some uninformative prior, or relative to some competing model(s), to measure how much information we gain by running the model (Nearing and Gupta, 2015). True science-focused efforts, however, where the intent is learning rather than prediction, cannot rely on this type of comparative hypothesis testing. We therefore encourage scientists to take the first perspective outlined above and to attempt to measure the ontological information that is lost by their models, rather than the epistemological information that is gained from their models. This represents a radical departure from how scientists usually approach the problem of model evaluation.

It turns out that it is possible to approximate the latter objective in practice. We are aware of no existing efforts to this effect in either the philosophy or practice of science (except by Gong et al., 2013, whose fundamental insight is the basis for this talk), and here I offer two examples of practical methods that scientists might use to approximately measure ontological information. I place this practical discussion in the context of several recent and high-profile experiments that have found that simple out-of-sample statistical models typically (vastly) outperform our most sophisticated terrestrial hydrology models. I offer some perspective on several open questions about how to use these findings to improve our models and understanding of these systems.

Cartwright, N. (1983) How the Laws of Physics Lie. New York, NY: Cambridge Univ Press.

Clark, M. P., Kavetski, D. and Fenicia, F. (2011) 'Pursuing the method of multiple working hypotheses for hydrological modeling', Water Resources Research, 47(9).

Cover, T. M. and Thomas, J. A. (1991) Elements of Information Theory. New York, NY: Wiley-Interscience.

Cox, R. T. (1946) 'Probability, frequency and reasonable expectation', American Journal of Physics, 14, pp. 1-13.

Csiszár, I. (1972) 'A Class of Measures of Informativity of Observation Channels', Periodica Mathematica Hungarica, 2(1), pp. 191-213.

Davies, P. C. W. (1990) 'Why is the physical world so comprehensible', Complexity, entropy and the physics of information, pp. 61-70.

Gong, W., Gupta, H. V., Yang, D., Sricharan, K. and Hero, A. O. (2013) 'Estimating Epistemic & Aleatory Uncertainties During Hydrologic Modeling: An Information Theoretic Approach', Water Resources Research, 49(4), pp. 2253-2273.

Jaynes, E. T. (2003) Probability Theory: The Logic of Science. New York, NY: Cambridge University Press.

Nearing, G. S. and Gupta, H. V. (2015) 'The quantity and quality of information in hydrologic models', Water Resources Research, 51(1), pp. 524-538.

Popper, K. R. (2002) The Logic of Scienti [U+FB01] c Discovery. New York: Routledge.

Van Horn, K. S. (2003) 'Constructing a logic of plausible inference: a guide to cox's theorem', International Journal of Approximate Reasoning, 34(1), pp. 3-24.