EGU24-20548, updated on 08 Oct 2024
https://doi.org/10.5194/egusphere-egu24-20548
EGU General Assembly 2024

Some Thoughts on Causal Inference, the Scientific Method, and Data Assimilation

Michael Ghil[1,2,3], Alberto Carrassi[4], and Olivier de Viron[5]
[1]Labo. de Météorologie Dynamique, Geosciences Dept., Ecole Normale Supérieure & PSL University, Paris, France (ghil@lmd.ipsl.fr)
[2]Atmospheric & Oceanic Sciences Dept., University of California at Los Angeles, Los Angeles, CA, USA (ghil@atmos.ucla.edu)
[3]Mathematics Dept., Imperial College London, London, United Kingdom (mghil@icl.ac.uk)
[4]Alberto Carrassi, Physics & Astronomy Dept., University of Bologna, Bologna, Italy (alberto.carrassi@unibo.it)
[5]Olivier de Viron, Littoral Environnement et Sociétés, UMR 7266, La Rochelle Université, La Rochelle, France (olivier.de_viron@univ-lr.fr)

Causal inference is at the heart of the scientific method as usually practiced. Still, Karl Popper (*The Logic of Scientific Discovery*, 1935/1959) tells us that a theory in the empirical sciences can never be proven: it can only be falsified, meaning that it can, and should, be scrutinized with decisive experiments. Even so, nobody that I know writes or publishes papers to disprove one's own theory, only an opposing theory. And the debate rages on.

At the heart of this session lies the question of whether, and how, one can prove, rather than just disprove, a causal link between phenomena in the empirical sciences. The session deals specifically with statistical, as opposed to dynamical methods. These methods have the advantage that they are essentially indifferent to any laws of, or other accumulated heuristic ideas on, the field to which they are being applied: whether the time series one considers are from the environmental sciences, biology or medicine does not matter, only their length and accuracy does.

Judea Pearl (e.g., *Stat. Surveys*, 2009) made an important observation on how to transcend the saying that "Correlation is not causation" by pointing out that standard methods of statistical analysis rely on the stationarity hypothesis of the phenomena being examined. Crucial questions, however, like the causal role of anthropogenic forcing in climate change, deal precisely with the causes of nonstationarity. In particular, Pearl suggested counterfactual analysis as an essential approach in establishing criteria for the necessary and sufficient character of a given cause for a given phenomenon. Thus, the common approach of detection and attribution in the climate sciences only covers the sufficiency aspect of anthropogenic forcing, and more can be done (Hannart et al., *BAMS*, 2016; *Clim. Change*, 2016).

The present talk will cover four specific aspects of these broad issues: (i) the distinction between information transfer, including both linear correlations and nonlinear extensions thereof, and true causation; (ii) the divergent results of some widely, and not so widely, used methods of studying information transfer (Krakovska et al., *PRE*, 2018; Kossakowski et al., *Psychol. Methods*, 2021;

Delforge et al., *HESS*, 2022); (iii) shared variability of climatic time series (De Viron, *GRL*, 2013;); and (iv) the uses of data assimilation in applying counterfactual theory to nonstationary phenomena (Carrassi, *QJRMS*, 2017; Metref et al., *QJRMS*, 2019).

Conclusions will include the obvious one that statistical studies of causal inference have to be complemented by dynamical ones.