

# Does influenza drive absolute humidity?

Edward B. Baskerville<sup>a</sup> and Sarah Cobey<sup>a,1</sup>

Multiple lines of evidence suggest that absolute humidity and school terms affect the timing of influenza epidemics, although the contributions of these and other factors are still unresolved (1–5). In PNAS, Deyle et al. (6) applied convergent cross-mapping (CCM) to measure the impact of environmental variables on influenza activity. The authors concluded that absolute humidity and temperature are not merely correlated with influenza activity but in fact directly or indirectly cause epidemics.

The mathematical basis of CCM involves strict assumptions that may be violated in nature (7). Process noise, secular changes in parameters, and other transient dynamics can lead to situations where the underlying attractor is undersampled, causing errors in reconstruction and inference (8). There are not yet any clear statistical tests to determine when CCM has gone awry.

One simple test of the method is whether it correctly predicts that influenza does not drive absolute humidity, temperature, relative humidity, and precipitation. We used the authors' code to test these interactions for each of the 26 countries shown in figure 3 of Deyle et al. (6), with modifications, described below, to support different test assumptions. Deyle et al. (6) used surrogate time series with randomized anomalies for each of the environmental variables to construct a null distribution for significance testing. Initially, we analogously used surrogate time series only for influenza to assess whether influenza affected each of these variables. In 53 of 104 tests influenza was inferred to drive the environmental variable. When using surrogate time series for both influenza and the environmental variable influenza was inferred

to drive the environmental variable in 56 of 104 tests (Fig. 1). This general result was robust to different transformations of the time series, including

- i) using surrogates for both influenza and the environmental variable, influenza alone, or the environmental variable alone;
- ii) using simple weekly means instead of fitted splines for the seasonal surrogates;
- iii) removing all zeros from the influenza time series and restricting analyses to periods when influenza was present;
- iv) log-transforming influenza incidence; and
- v) additionally requiring that the maximum cross-correlation appear at a lag corresponding to the correct directionality in time (9).

Summaries of results from different test variants are in Table 1. More thorough testing with simulated data shows that using CCM with surrogate time series frequently leads to false positives. Other criteria for causality perform better in some instances, but none is universally reliable in a seasonally forced infectious disease system (8).

Although it may be possible to choose surrogates and transformations to obtain more favorable results, it is clear that more work is needed on the statistical basis of CCM before it can be applied out of the box. We agree with Deyle et al. (6) that weather affects the timing of influenza activity, but we do not think that analysis based on CCM supports the correct inference regarding causal direction.

The code for our analysis, along with source data and full results, is available at [github.com/cobeylab/ccm-letter](https://github.com/cobeylab/ccm-letter).

<sup>a</sup>Department of Ecology and Evolution, University of Chicago, Chicago, IL 60637

Author contributions: E.B.B. and S.C. designed research; E.B.B. performed research; E.B.B. analyzed data; and E.B.B. and S.C. wrote the paper. The authors declare no conflict of interest.

<sup>1</sup>To whom correspondence should be addressed. Email: [cobey@uchicago.edu](mailto:cobey@uchicago.edu).

