LETTER

Questionable dynamical evidence for causality between galactic cosmic rays and interannual variation in global temperature

Tsonis et al. (1) claim a significant causality between cosmic rays (CRs) and interannual variation in global temperature (ΔGT) by using convergent cross mapping (CCM) (2). Their results are potentially helpful to model the climate system. However, we are unable to reproduce their results and suspect that this dynamical evidence uncovered by CCM is questionable.

CCM is based on the theory of state space reconstruction, and it was originally introduced by Rulkov et al. (3). If variable Y unidirectionally causes X, then the values of Y can be reconstructed from the state information of X via cross mapping (denoting by X xmap Y) based on a library (=training dataset) $M_{X,Y}$ containing pairs of $(\vec{x}, y)$ recorded simultaneously, but not vice versa. For causality, the prediction performance of X xmap Y improves with increasing library size L, which is what convergence means.

Choosing libraries suitably is crucial. Different methods may lead to different results. However, there is no evidence in refs. 1, 2, and 4 that the authors are aware of this problem. No relevant information can be found in ref. 1. Different methods are used in refs. 2 and 4. According to the shared code in ref. 4, they first form $\widetilde{M}_{X,Y}$ containing all pairs $(\vec{x}, y)$ and then, for a given L, construct $M_{L,X,Y}$ by sampling pairs from $\widetilde{M}_{X,Y}$ using the bootstrap method (method 1). In ref. 2, $\vec{x}$ constructed from a segment of X and the corresponding y are used to form $M_{L,X,Y}$ (method 2). For unbiased results, it is essential that, for each predictee y_i, the pair $(\vec{x}_i, y_i)$ must be excluded from $M_{L,X,Y}$. Thus, the library size after exclusion may shrink for methods 1 and 2. Therefore, we propose method 3 which constructs libraries by bootstrapping from $\widetilde{M}_{X,Y} \setminus \{(\vec{x}_i, y_i)\}$ for each predictee y_i. Advantages of this method over methods 1 and 2 are that it keeps the library size fixed and removes any nonstationarity of libraries reconstructed on the basis of a segment of X by uniformly sampling from the whole embedding space. Therefore, method 3 is recommended when implementing CCM.

Using methods 1–3, results in Fig. 1 show no evidence for causality between CR and ΔGT. Studying the prediction performance for many L values is a version of multiple testing (5). A statistically significant rejection of the null hypothesis requires that significantly more than 5% of all trials lie outside the 95% confidence interval, which is not given if a single value is outside (as in ref. 1). Alternatively, one could enlarge the confidence interval using, e.g., a Bonferroni correction (which is not necessary in Fig. 1, because all curves are located inside the confidence intervals of null hypotheses).

In summary, our analysis of the same datasets does not yield any statistically significant indication. Considering the possible effect of multiple testing, it is more appropriate to conclude that no significant causality between CR and ΔGT can be uncovered by CCM.

Ming Luo[a], *Holger Kantz*[b], *Ngar-Cheung Lau*[a], *Wenwen Huang*[b], *and Yu Zhou*[b,1]

[a]*Institute of Environment, Energy and Sustainability, The Chinese University of Hong Kong, Shatin, NT, Hong Kong; and* [b]*Max Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany*

1 Tsonis AA, et al. (2015) Dynamical evidence for causality between galactic cosmic rays and interannual variation in global temperature. *Proc Natl Acad Sci USA* 112(11):3253–3256.

2 Sugihara G, et al. (2012) Detecting causality in complex ecosystems. *Science* 338(6106):496–500.

3 Rulkov NF, Sushchik MM, Tsimring LS, Abarbanel HDI (1995) Generalized synchronization of chaos in directionally coupled chaotic systems. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* 51(2):980–994.

4 Clark AT, et al. (2015) Spatial convergent cross mapping to detect causal relationships from short time series. *Ecology* 96(5): 1174–1181.

5 Miller RG (1981) *Simultaneous Statistical Inference* (Springer, New York).

Author contributions: H.K. and Y.Z. designed research; M.L., H.K., and Y.Z. performed research; H.K. contributed new reagents/analytic tools; M.L., W.H., and Y.Z. analyzed data; and M.L., H.K., N.-C.L., and Y.Z. wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. Email: zhouyu@pks.mpg.de.

 www.pnas.org/cgi/doi/10.1073/pnas.1510571112

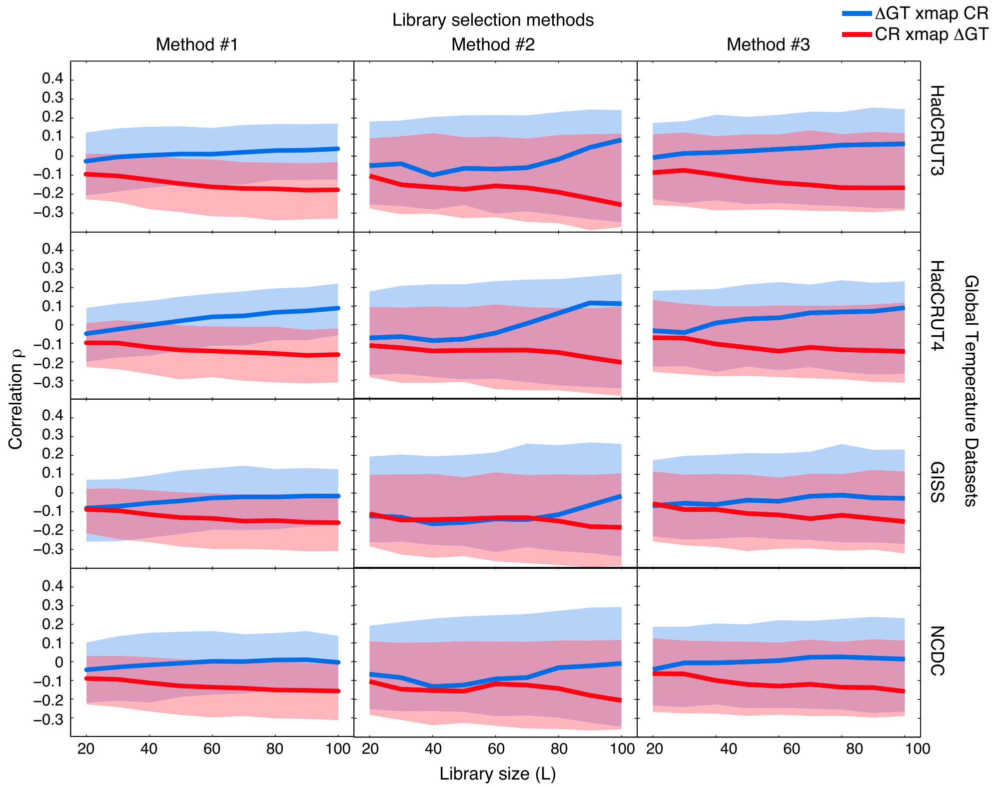

Fig. 1. Correlation (ρ) of cross-mapped versus observed values as a function of the size (L) of the library selected by three different methods (by column) applied to four different global temperature datasets (by row). Blue and red curves indicate ΔGT xmap CR (i.e., CR causes ΔGT) and CR xmap ΔGT (i.e., ΔGT causes CR), respectively. Shaded areas with the corresponding colors are the 95% confidence intervals of null hypotheses. Results (curves and confidence intervals) shown in the first column are obtained by using the R package shared in ref. 4. Confidence intervals in the second and third columns are determined by 1,000 phase-randomized surrogate time series (the method used in ref. 1). If one curve is located inside the corresponding shading, then there is no significant causality at the 0.05 level. To make a fair comparison, the embedding time lag, embedding dimension, and optimal time lags are set to the same as those in ref. 1. Four global temperature datasets are HadCRUT3 (the one used in ref. 1), HadCRUT4, Goddard Institute for Space Studies (GISS), and NOAA National Climatic Data Center (NCDC) datasets. Although different selection methods lead to different results, none suggests significant causality between CR and ΔGT. In addition, we have done the same analysis but without time shift between CR and ΔGT; and for method 1 we have also used the phase-randomized surrogate to determine the confidence interval. In these cases, results still do not show any significant causality. Our conclusion is robust to library selection methods and temperature datasets.