

Methods note · Sea ice and ocean

Look for a common driver before linking ocean series

Amina Okafor · 16 October 2026

Fictional author and institute · Editorial demonstration, not peer reviewed

ABSTRACT

Strong correlation and a plausible lead can arise when two ocean records share a driver rather than directly influence one another. This methods note combines Earth-system causal-inference research with studies of nonlinear ecological dynamics and autocorrelated fish-environment data. An invented two-regime calculation produces a correlation of approximately 0.94 despite zero within-regime association and no constructed direct effect. The analysis then examines lagged common forcing, sampling and smoothing, seasonal transformations, adjustment choices, and the assumptions behind advanced causal methods. It attributes no actual ocean event and does not claim that conditioning alone establishes causation. Its practical outcome is a sequence of rival physical explanations and discriminating observations, allowing unresolved associations to guide a better next measurement instead of an overstated causal story.

Keywords: common drivers · ocean time series · causal inference · autocorrelation · lagged association

A source-based editorial text, not a report of original Still North research. Examples and imagery are illustrative.

Two ocean records rise together, and one leads the other by a month. That pattern can motivate a mechanism, but it does not establish that the first variable caused the second. A seasonal cycle, wind regime, moving water mass, or shared processing step can generate association without a direct link. Before choosing an impressive causal method, draw the simpler possibility that both records respond to something else.

This methods note uses published work on Earth-system causal inference, nonlinear ecological dynamics, and autocorrelated fish-environment records to develop that first comparison. It offers a hypothetical common-driver calculation and a practical analysis sequence. It does not attribute a real ocean event, apply a causal-discovery algorithm, or claim that statistical adjustment alone turns observational records into an experiment.

Write down the rival explanations

Suppose salinity at mooring A and salinity at mooring B covary. One hypothesis is transport: changes near A move downstream to B. Another is a common wind regime that changes mixing or circulation at both sites. A third is that both instruments experience the same seasonal water-mass transition. A fourth is a shared processing artefact. These hypotheses can produce similar plots while making different predictions about timing, currents, vertical structure, and responses during unusual events.

Runge and colleagues explain how common drivers and indirect pathways complicate correlation-based interpretation in Earth-system science. Their examples distin-

guish pairwise association from a multivariable account of the underlying relationships. The lesson is not that one algorithm automatically discovers the truth. Different causal frameworks require different assumptions about observed drivers, time order, functional relationships, and the dynamics represented in the record.

A worked common-driver example

Consider eight invented cases split between two wind regimes. The numbers below are arbitrary index values, not salinity measurements. In the low regime, A takes values 0, 0, 1, and 1, while B takes 0, 1, 0, and 1. In the high regime, add four to every value in both series. Across all cases, A and B have a strong positive association because the regime shifts both records upwards.

Using all eight cases, each series has mean 2.5 and population variance 4.25. Their covariance is 4, giving a correlation of $4/4.25$, approximately 0.94. Within each regime, the covariance is zero. The strong aggregate relationship was constructed entirely by a shared regime shift; no direct influence between A and B was included. This is a teaching demonstration, not a claim that conditioning on a wind index will remove real ocean associations.

In actual records, the driver may be measured imperfectly, act with different lags, or be missing altogether. Residual association after adjustment therefore does not prove a direct link, and disappearance after adjustment does not establish that the driver was the correct causal explanation. The calculation identifies one rival mechanism that a pairwise correlation cannot distinguish. Phys-

Hypothetical regime	A values	B values	Within-regime correlation
Low	0, 0, 1, 1	0, 1, 0, 1	0
High	4, 4, 5, 5	4, 5, 4, 5	0

ical evidence must decide whether the proposed regime is plausible.

A lead does not eliminate the common cause

A shared forcing can affect two sites at different speeds. Wind might change a shallow mixed layer promptly and a deeper layer later, or currents can redistribute a water mass along different pathways. One observed series then leads another even without a direct effect from the first measured variable to the second. A lead-lag plot should therefore be compared with plausible forcing and transport timescales, not interpreted as a directional arrow by itself.

Sampling and processing can also generate timing differences. Monthly means, interpolation, smoothing, and unequal observation intervals alter phase and persistence. A sensor moving vertically can encounter a property gradient that resembles temporal change. Before testing a mechanism, reconcile clocks, depths, spatial support, missingness, aggregation, and filter choices. Two carefully calibrated instruments still produce incomparable series if their measurement frames answer different physical questions.

Autocorrelation changes the evidence, not just a formula

Ocean records often retain memory. A warm month can be followed by another warm month without providing a new independent realisation of the same question. Counting every day as independent can exaggerate the evidence for a correlation. Pyper and Peterman compared methods for autocorrelated fish and environmental records using simulations, showing that accounting for dependence and removing it are not interchangeable procedures.

In particular, first-differencing or prewhitening can remove slow variation that is itself scientifically important. Their study found that this can reduce power in some circumstances, while helping in others. The appropriate choice depends on the timescale of the proposed relationship. A blanket instruction to detrend or difference every ocean record can destroy the signal under investigation just as surely as ignoring dependence can overstate confidence.

Seasonal anomalies should use a documented reference and comparable observation coverage. Removing the average annual cycle asks whether departures covary, not whether the full seasonal mechanism exists. Detrending asks another conditional question. Show the original and transformed records, explain which component is being tested, and report whether conclusions depend on that choice. Avoid selecting the transformation that produces the smallest probability value after inspecting the outcome.

Condition on causes thoughtfully

More covariates do not automatically make an analysis causal. A common cause is a candidate adjustment variable, but a mediator may carry the very effect being estimated. Conditioning on it changes the question from total effect to something closer to a direct effect under additional assumptions. A variable influenced by both series can introduce association when conditioned on.

The causal diagram should determine the proposed adjustment set, not a menu of all available columns.

Physical process knowledge helps here. If transport is proposed, seek current measurements, plausible pathways, water-mass properties, and a travel time consistent with the lag. If shared wind forcing is proposed, examine the forcing at the appropriate spatial and temporal scale rather than substituting an arbitrary distant climate index. New observations should discriminate among explanations, not merely provide a third line that happens to resemble the first two.

What advanced methods add, and what they do not

Sugihara and colleagues developed convergent cross mapping for causal relationships in complex dynamical systems. That is a different framework from a simple regression and has requirements about the dynamics recoverable from the observed series. Its existence does not mean that every pair of short, noisy, seasonally forced records contains enough information for reliable directional inference. The method's assumptions must match the system and sampling.

Runge and colleagues' nonlinear time-series method addresses causal associations in larger multivariable datasets, including challenges from autocorrelation. Such methods can make assumptions explicit and test conditional relationships more appropriately than unrestricted pairwise searches. They still cannot guarantee that every relevant common driver was observed, that the system remained stationary, or that time aggregation preserved causal ordering. A sophisticated network diagram is a hypothesis-bearing result, not a photograph of causation.

A proportionate analysis and its limits

Begin with the physical question and rival diagrams. Audit the measurement chain. Plot relevant timescales and event timing. Define transformations and candidate lags from mechanisms where possible, then use an inference method suited to the dependence and observation process. Report sensitivity to plausible drivers and model choices, and account for multiple searches if many variables or lags were explored. Hold back a period for predictive checking where the record permits, while remembering that successful prediction alone does not establish causation.

This sequence cannot replace interventions, complete observing systems, or independent mechanistic evidence. Short records can leave the rival explanations unresolved. That is a useful outcome if stated precisely: the series covary at a specified scale, and the available data cannot distinguish direct transport from shared forcing. The next measurement then has a purpose. Looking for a common driver is not an obstacle to explanation; it is how an appealing two-line story becomes a scientific question.

References

1. Runge, J., et al. (2019). Inferring causation from time series in Earth system sciences. *Nature Communications*, 10, 2553.
2. Pyper, B. J., and Peterman, R. M. (1998). Comparison of methods to account for autocorrelation in corre-

- lation analyses of fish data. *Canadian Journal of Fisheries and Aquatic Sciences*, 55, 2127–2140.
3. Sugihara, G., et al. (2012). Detecting causality in complex ecosystems. *Science*, 338, 496–500.
 4. Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., and Sejdinovic, D. (2019). Detecting and quantifying causal associations in large nonlinear time series datasets. *Science Advances*, 5, eaau4996.