## **Box 3: Reality Bites: Coping with spatiotemporal omitted confounders**

Spatiotemporal confounding variables – those that are site (or plot) specific and vary through time – pose challenges, and the solutions require more thoughtful study and statistical model design. To illustrate, we consider a scenario where recruitment, a confounding variable related to both snail abundance and temperature, is not static through time but instead varies by site and year (as in a realistic case). For example, sites that experience strong cold-water pulses in a year also experience unusually snail high recruitment in those same years due to oceanographic drivers. The sampling designs for coping with spatio-temporal omitted variables are based on the same principles as cross-sectional and longitudinal sampling, only now we combine the two.

With longitudinal data that includes multiple plots sampled within a site through time, we can flexibly control for this sort of spatiotemporal confounding at the site level by extending the two-way fixed effect designs discussed above. We can add a site-by-time fixed effect, , to our model, in addition to a fixed effect of plot, , where k is a fixed plot within site resampled over time (see below for a discussion of fixed versus re-randomized plots). This produces the following means model:

From this equation, we can see that captures time invariant plot-level confounding effects while captures the effects of spatiotemporal omitted variables at the site by time level. Note, there could be additional spatial or temporal only confounders. This model design sweeps their effects onto the spatiotemporal term.

In small datasets, the above model design can consume degrees of freedom rapidly. For this reason, we can instead use the more efficient correlated random effects model design (e.g., a variation on the Two-way Mundlak model design sensu Wooldridge 2021) using site-year means () and plot means () for the entire survey to control for spatiotemporal and plot confounding respectively:

Here the and terms are random effects for plot and unique site-time combinations respectively.

When sampling to handle spatiotemporal confounders, should plots within sites over time be permanent or randomly placed each year? The above models assume permanent plots, so we can eliminate confounding variables at the plot-level that is time invariant over the study period. For this reason, permanent plots help us cope with within-site OVB issues and have higher power to detect change over time (Urquhart & Kincaid 1999). Logistically, however, permanent plots within sites might not be possible. As such, the above models can be modified to drop plot effects; however, they would then assume that there are no confounding differences across plots and could have lower power to detect effects of drivers. We emphasize that the choice of fixed or random plot placement with these designs is a balancing act, however, as fixed plots can lead to a lower sample size due to logistical considerations in many environments, and direct readers to other explorations of this topic (see Gomes 2022 for an excellent jumping off point). Finally, without a nested data structure – e.g., plots within sites resampled over years – we cannot include a site by year effect as in the above models. We can attempt to use site-level polynomials or GAMs to mimic site-by-time effects (for a polynomial example see Dee *et al.* 2016), but, this requires knowledge of how the confounder varies at sites over time and extensive testing for robustness to these assumptions. In the many cases where this is not possible or inadvisable given the likelihood of creating incorrect causal inference, without multiple plots per site over time, “nothing to be done” (Beckett 1954).

In general, we urge caution when dealing with spatiotemporal omitted variables, and careful use of causal diagrams to ensure that we are controlling for a confounder at the relevant spatiotemporal scale. This topic is one that that deserves far more exploration in Ecology. More from other disciplines on this tricky class of problem and approaches can be found from literature outside of the scope of this paper (e.g., Ferraro & Hanauer 2014; Athey & Imbens 2017; Oster 2019).