## **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 can 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 joint drivers of oceanography. The sampling designs for coping with spatio-temporal omitted variables are based on the same principles as before, only now requiring a multiple plots per site per year, as the spatiotemporal confounders do not vary at the plot scale.

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 such that we do not have to estimate additional parameters.

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 spationtemporal and plot confounding respectively:

Here the terms are random effects for plot and unique site-time combinations. Some of these could be unnecessary depending on relevant sources of confounding variation (e.g., perhaps only site-time is necessary). If a researcher is interested in estimating additional terms looking at spatially confounded gradients or temporally confounded trends, they could add those trends using the appropriate site- or year-level averages and random effects.

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. Permanent plots allow for plot-level effects which can cope with within-site OVB issues and have higher power to detect change over time (Urquhart & Kincaid 1999). Logistically, permanent plots might not be possible. As such, the above models can be modified to have no plot effects, as we assume that re-randomization removes confounding due to repeated sampling of the same plots. The resulting analyses should perform, although the estimate of the temperature effect will not be as precise due to greater residual error. If a researcher is worried about spatial confounding due to plots within a site being similar, they can include a site-level effect (either fixed or group mean covariate and random effect). We emphasize that it 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 still have some options, however, although they can be more *ad hoc*. See supplementary materials S2.

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).