Notation

$\mathbf{Y} \equiv$ One-dimensional vector containing the outcome measure of interest for each unit of observation; i.e., a cross-sectional measurement of forest cover change. Y_i represents the outcome measurement at unit of observation i.

$\mathbf{X} \equiv$ j by i matrix containing ancillary information which may impact the outcome measure of interest, excluding the treatment. $X_{j,i}$ represents the information for covariate j at unit of observation i.

$\mathbf{T} \equiv$ One-dimensional vector containing the treatment status for each unit of observation; i.e., if a project to decrease deforestation existed at that location. T_i represents the treatment status at unit of observation i.

Here, we examine the impact of spatial spillover in any of the three elements **Y,X,T** in matching methods for causal inference. Spatial spillover can be simply understood as the presence of correlation between measurements in neighboring units. For example:

1. **Y** - in the case of deforestation, communities that practice lumber harvesting as a livelihoods strategy may suggest that to neighboring communities, or even enter into neighboring lands to conduct lumber operations.

2. **X** - population could very reasonably be inferred to be a driver of deforestation; neighboring communities may have similar populations due to a passed on cultural demographic characteristic.

3. **T** - a treatment - i.e., paying a community to prevent deforestation - could spread to neighboring communities due to word-of-mouth, or treatment designs to drive down costs by reducing travel times.

These are just illustrative examples, but highlight the potential challenge that modeling spillover effects can cause.

X

$$ \tag{1} $$