Stream network models

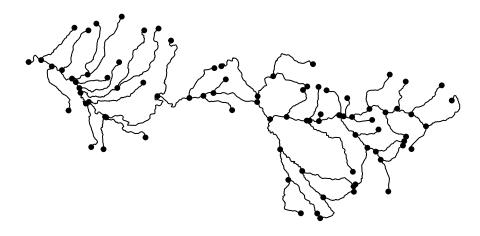
James T. Thorson

```
library(sf)
library(sfnetworks)
library(tinyVAST)
library(viridisLite)
set.seed(101)
```

tinyVAST is an R package for fitting vector autoregressive spatio-temporal (VAST) models using a minimal and user-friendly interface. We here show how it can fit a stream network model, where spatial correlations arise from stream distances along a network.

First, we load a shapefile representing a stream network, and convert it to *sfnetwork* format. This format includes edges representing stream segments, and nodes where edges connect.

East Fork Lewis Basin



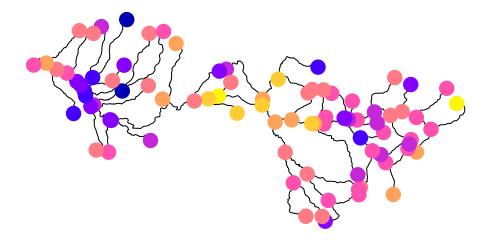
We then convert it to an S3 class "sfnetwork_mesh" defined by tinyVAST for stream networks, and rescale distances to 1000 ft (to ensure that distances are 0.01 to 100, avoiding issues of numerical under or overflow).

```
# Rescale
graph = sfnetwork_mesh( stream )
graph$table$dist = graph$table$dist / 1000 # Convert distance scale
```

Next, we'll simulate a Gaussian Markov random field at stream vertices using simulate_sfnetwork, sample evenly spaced locations along the stream using st_line_sample, project the GMRF to those locations using sfnetwork_evaluator, and simulate data at those locations:

```
omega_s = simulate_sfnetwork( n=1, sfnetwork_mesh=graph, theta=kappa)[,1]
# sample locations along network
extrap = st_union( st_line_sample( activate(stream, "edges"), density=1/10000))
extrap = st_cast( extrap, "POINT" )
# Project to sampled locations
A_is = sfnetwork_evaluator( stream = graph$stream,
                               loc = st_coordinates(extrap) )
omega_i = (A_is %*% omega_s)[,1]
# Simulate sampling
#Count = rpois( n=graph$n, lambda=exp(alpha + omega) )
Count_i = rnorm( n=length(omega_i), mean=alpha + omega_i, sd=0.5 )
\# Format into long-form data frame expected by tiny VAST
Data = data.frame( Count = Count_i,
                  st_coordinates(extrap),
                   var = "species", # Univariate model so only one value
                   time = "2020", \# no time-dynamics, so only one value
                   dist = "obs" )  # only one type of sampling in data
```

We can visualize the GMRF at those locations using $\mathit{sfnetwork}$



Finally, we can fit the model, interpolate the GMRF along at dense locations along the stream network, and plot those with the true (simulated) values at the location of simulated samples.

omegahat_i

