Alternative hypothesis testing to study eco-evolutionary dynamics

Jelena H. Pantel

2022-09-12

```
library(ecoevor)
library(deSolve)
library(tidyverse)
library(ggplot2)
library(gen3sis)
library(raster)
library(MASS)
library(Matrix)
library(graphics)
library(stringr)
library(grDevices)
library(stats)
library(here)
library(rmarkdown)
```

1 Overview

Eco-evolutionary hypothesis testing is implemented by using theoretical or simulation models to compare alternative mechanistic processes that may be structuring data. We introduce two eco-evolutionary models and for each, we run simulations under parameter values or model components that match distinct eco-evolutionary hyotheses. In a separate vignette *ecoevo-ABC*, we describe how to use Approximate Bayesian Computing (and alternatively machine learning) to compare data generated under these alternative hypotheses to observed data.

2 Example 1. pathogen-virulence model

In Figure 1 of Pantel & Becks (in review), we discuss a host-pathogen model of eco-evolutionary dynamics described in Luo & Koelle (2013). The model equations are given as:

$$\frac{dS}{dt} = B - \beta(\alpha)SI - \mu S$$

$$\frac{dI}{dt} = \beta(\alpha)SI - (\mu + \alpha)I$$

$$\frac{d\alpha}{dt} = V(S\frac{d\beta}{d\alpha} - 1) = V(S\frac{c}{\gamma}\alpha^{(\frac{1}{\gamma})-1} - 1)$$

where S and I are the susceptible and infected host population size, B is a constant birth rate, μ is the per capita death rate, α is the disease-induced mortality rate (virulence), and β is the transmission rate. The

transmission-virulence trade-off is modeled by $\beta(\alpha) = c\alpha^{\frac{1}{\gamma}}$, where c is a positive constant, γ indicates the magnitude of the transmission-virulence tradeoff, and V is the additive genetic variance.

Virulence (α) can evolve if the additive genetic variance for this trait $V > 0$, and the rate of change for this trait depends on the susceptible host population size $(d\alpha/dt \sim f(S))$. This link between the evolving trait and host population dynamics can represent an eco-evolutionary feedback as $S \sim f(\alpha)$ and $d\alpha/dt \sim f(S)$ (Lion 2018).

In such a system, ecological, evolutionary, and interacting eco-evolutionary processes may be occurring. A researcher can generate simulation models under a series of alternative hypotheses and compare simulated to observed data. We consider population dynamics for (1) Hypothesis 1: virulence evolution is not possible (the additive genetic variance for virulence $V = 0$), (2) Hypothesis 2: virulence evolution is possible ($V > 0$) and there are eco-evolutionary feedbacks ($S \sim f(\alpha)$ and $d\alpha/dt \sim f(S)$), and (3) Hypothesis 3: virulence evolution is possible ($V > 0$) but eco-evolutionary dynamics are removed from the system by decoupling the dependence of virulence evolution from host population dynamics ($d\alpha/dt \,!\sim f(S)$). The population dynamics for all three hypothesized scenarios are shown in Figure 1.

```
# H1. No evolution (V = 0)
parms <- c(B = (10^7), mu = 1, V = 0, gam = 2, c = 3 * 10^-7)
times <- seq(from = 0, to = 40, by = 1)
xstart <- c(S = 10^7, I = 1, alpha = 4)
h1 <- hostpath_evol(times, xstart, parms, method = "ecoXevo")


# H2. Evolution (V = 4)
parms <- c(B = (10^7), mu = 1, V = 4, gam = 2, c = 3 * 10^-7)
times <- seq(from = 0, to = 40, by = 1)
xstart <- c(S = 10^7, I = 1, alpha = 4)
h2 <- hostpath_evol(times, xstart, parms, method = "ecoXevo")


# H3. Evolution (V = 4), no eco-evo feedback
parms <- c(B = (10^7), mu = 1, V = 5, gam = 2, c = 3 * 10^-7)
times <- seq(from = 0, to = 40, by = 1)
xstart <- c(S = 10^7, I = 1, alpha = 4)
h3 <- hostpath_evol(times, xstart, parms, method = "eco_evo")
```
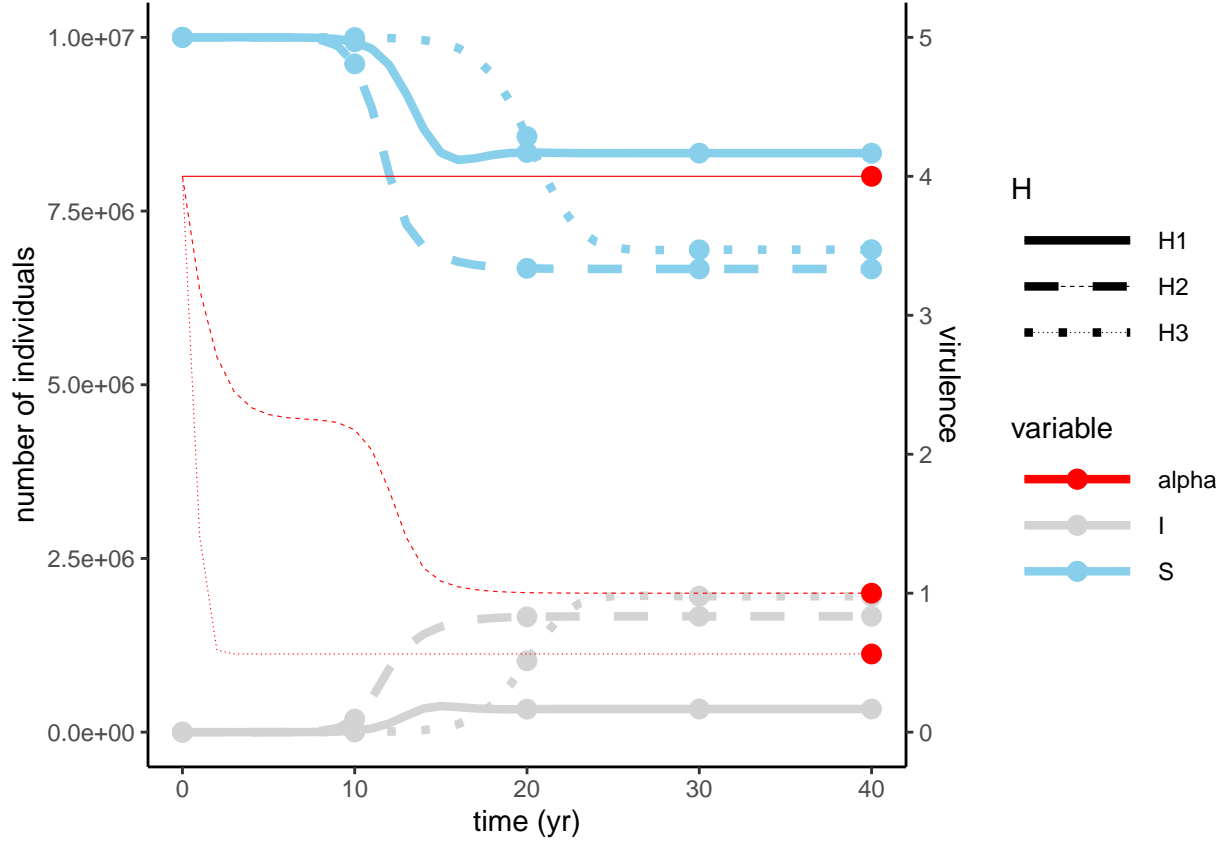
3 Example 2. macroevolution, speciation, and community assembly

In Figure 2 of Pantel & Becks (in review), we discuss a model of evolution, speciation, and community assembly using the gen3sis eco-evolutionary simulation engine (Hagen et al. 2021). We adapt their model with a hypothetical community of 10 ancestral species, randomly placed as a single initial population on the map of South America (an 81x71 grid with cell-specific temperature values that impact population growth via the distance between a local population's trait value and the local optimum and that impact local carrying capacity). Movement is via a cost function based on the distance between sites and a species-specific dispersal value, and evolution and speciation occurs across the dynamic landscape over 40 million years.

In such a system, ecological and evolutionary processes combine to structure biodiversity patterns at multiple scales, but determining the nature of which processes are most likely to drive observed contemporary biodiversity data can be difficult. However, a researcher can generate simulation models under a series of alternative hypotheses and compare simulated to observed data. We consider eco-macroevolutionary dynamics for:

(1) Hypothesis 1 (H1): evolution in a single trait - each deme experiences evolution in the temperature trait (that determines the population growth), with mean traits of the population after evolution $x_{\mu-temp}$ drawn from a random normal distribution centered on parental trait values x_{temp} and mutation standard deviation of $\sigma_{\mu-temp} = 0.001$.

(2) Hypothesis 2 (H2): evolution in two traits that evolve at the same rate - evolution in an additional trait, dispersal (which determines the connectivity between sites and colonization of new sites) is considered, with $x_{\mu-temp} \sim \mathrm{N}(x_{disp}, \sigma_{\mu-disp} = 0.001)$.

(3) Hypothesis 3 (H3): evolution in two traits that evolve at differing rates - evolution of the dispersal trait is much faster than the temperature trait ($\sigma_{\mu-temp} = 0.001$, $\sigma_{\mu-disp}$).

(4) Hypothesis 4 (H4): evolution in two correlated traits - evolution of the dispersal trait is much faster than the temperature trait (as in e) and the traits are correlated ($r_{temp-disp}$ = -0.25).

For all four hypotheses, the final alpha diversity (local species richness) and time series of some summary statistics (number of speciation and extinction events, total regional richness) are shown. These summary statistics create features that could be used for comparison to observed data to distinguish among the alternative hypotheses that may have structured observed data.

```r
datapath <- here::here("inst", "extdata", "Example2")
## Create environmental rasters bricks
temperature_brick <- brick(paste(datapath, "/input_rasters/temperature_rasters.grd",
    sep = ""))
aridity_brick <- brick(paste(datapath, "/input_rasters/aridity_rasters.grd",
    sep = ""))
area_brick <- brick(paste(datapath, "/input_rasters/area_rasters.grd",
    sep = ""))
## Create list with environmental raster bricks
landscapes_list <- list(temp = NULL, arid = NULL, area = NULL)
for (i in 1:nlayers(temperature_brick)) {
    landscapes_list$temp <- c(landscapes_list$temp, temperature_brick[[i]])
    landscapes_list$arid <- c(landscapes_list$arid, aridity_brick[[i]])
    landscapes_list$area <- c(landscapes_list$area, area_brick[[i]])
}
## Create cost function
cost_function_water <- function(source, habitable_src, dest,
    habitable_dest) {
    if (!all(habitable_src, habitable_dest)) {
        return(2/1000)
    } else {
        return(1/1000)
    }
}
## Create input landscape files
gen3sis::create_input_landscape(landscapes = landscapes_list,
    cost_function = cost_function_water, directions = 8, output_directory = file.path(datapath,
        "created_landscapes"), timesteps = paste0(seq(65, 0,
        by = -1), "Ma"), crs = "+proj=longlat +ellps=WGS84 +datum=WGS84 +no_defs",
    calculate_full_distance_matrices = F)

## Hypothesis 1: Temperature trait evolves
sim1 <- gen3sis::run_simulation(config = file.path(datapath,
    "config/config_southamerica_H1.R"), landscape = file.path(datapath,
    "created_landscapes"), output_directory = file.path(datapath,
    "output"), call_observer = 1, verbose = 0)

## Hypothesis 2: Temperature and dispersal traits evolve,
## same speed
sim2 <- gen3sis::run_simulation(config = file.path(datapath,
    "config/config_southamerica_H2.R"), landscape = file.path(datapath,
    "created_landscapes"), output_directory = file.path(datapath,
    "output"), call_observer = 1, verbose = 0)

## Hypothesis 3: Temperature and dispersal traits evolve,
## dispersal trait evolves faster than temperature trait
```

```
sim3 <- gen3sis::run_simulation(config = file.path(datapath,
    "config/config_southamerica_H3.R"), landscape = file.path(datapath,
    "created_landscapes"), output_directory = file.path(datapath,
    "output"), call_observer = 1, verbose = 0)

## Hypothesis 4: Temperature and dispersal traits evolve,
## dispersal trait evolves faster than temperature trait,
## traits correlated 0.25
sim4 <- gen3sis::run_simulation(config = file.path(datapath,
    "config/config_southamerica_H4.R"), landscape = file.path(datapath,
    "created_landscapes"), output_directory = file.path(datapath,
    "output"), call_observer = 1, verbose = 0)
```
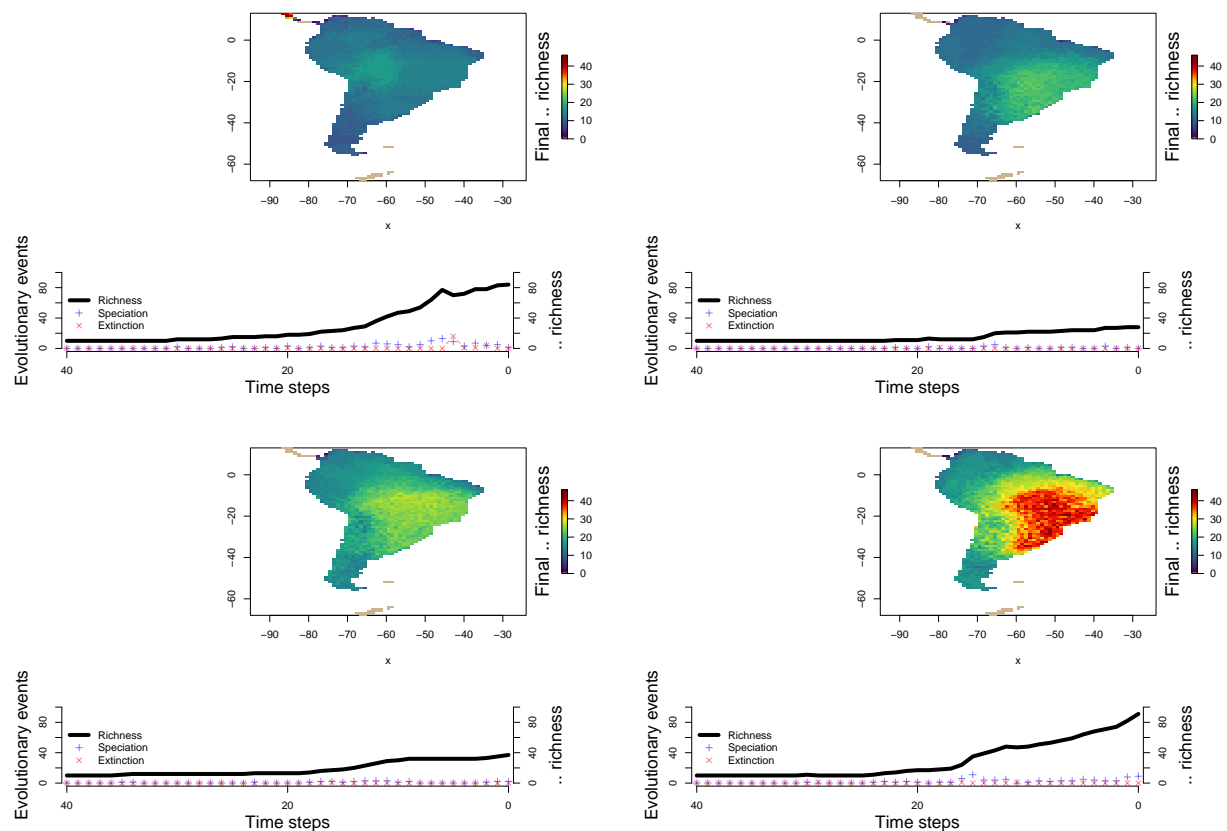

References

Hagen, O., Flück, B., Fopp, F., Cabral, J. S., Hartig, F., Pontarp, M., Rangel, T. F.,& Pellissier, L. (2021). gen3sis: A general engine for eco-evolutionary simulations of the processes that shape Earth's biodiversity. PLoS biology, 19(7), e3001340.

Lion, S. (2018). Theoretical approaches in evolutionary ecology: environmental feedback as a unifying perspective. The American Naturalist, 191(1), 21-44.

Luo, S., & Koelle, K. (2013). Navigating the devious course of evolution: the importance of mechanistic models for identifying eco-evolutionary dynamics in nature. The American Naturalist, 181(S1), S58-S75.