- n_s is the number of sites,
- N_{ts} be the number of species at time t and site s,
- d_{si} be the distance between two sites s and site i,
- The environmental variables are e_{chi}, e_{sst} and e_{upw},
- Let $y_{ts} = \log N_{(t+1)s}$.

The raw model is:

$$y = K \sum_{i=1}^{n_s} N_{ti} e^{-\frac{(d_{si}-\mu_d)^2}{\sigma_d^2}} + \beta_1 e_{chi} + \beta_2 e_{sst} + \beta_3 e_{upw} + \epsilon, \tag{1}$$

where $\epsilon \sim N(0, \sigma^2)$. So we have

- three parameters related to the **dispersal**: K, μ_d, σ_d.
- three parameters for the **environmental** variables $\beta_1, \beta_2, \beta_3$.

Let $\theta = (K, \mu_d, \sigma_d, \beta_1, \beta_2, \beta_3)$, x be our data (N:number of species, d:distance between the sites, e: the environmentaall variables), $y = f(x, \theta)$ is our model. Then $y \sim N(f(x, \theta), \sigma^2)$

$$L(\theta; y, x) = \prod_{j=1}^{m} \phi(\frac{y_j - f(x_j, \theta)}{\sigma}) \tag{2}$$

and

$$l(\theta; y, x) = \log L(\theta; y, x) = \sum_{j=1}^{m} \log \phi(\frac{y_j - f(x_j, \theta)}{\sigma}) \tag{3}$$

As for our data, we have 4 years(year0, year1, year3, year4) data. So y should be

y=N_{ts} for [year1(s1,...,s48),year2(s1,...,s48),year3(s1,...,s48)]

Predictive variables x contains three parts:
1. species number N,

N=N_{ts} for [year0(s1,...,s48),year1(s1,...,s48),year2(s1,...,s48)]

2. the distance d, which should be a distance matrix between the sites (s1,...,s48)
3. the environmental variables e_i, which is

e_i=e_i for [year0(s1,...,s48),year1(s1,...,s48),year2(s1,...,s48)]

Need to do: put the data in the log likelihood, and find the argmax of the parameters.