Affinity

Helmut Hillebrand

10 5 2022

Approach

The idea to use affinity as a measure of turnover arose, because we use dissimilarity metrics that reach a specific maximum (1 being the ultimate limit, but each system asymptotically reaching a max turnover) over time. This is similar to enzyme kinetics in response to resource availability, which uses Michaelis-Menten (MM) equations to estimate this maximum (max), but also the half-saturation concentration (ks), i.e. the amount of resource to achieve half-maximal kinetics. It has been proposed and tested that the ratio "max/ks" is a very good descriptor of the affinity of the population to this resource, as it reflects the ability to capture resources at very low concentrations. Transferring this to the turnover question, ks would be the halving distance (HD, the time needed to reduce initial similary by 50%) and affinitiy = max/HD would be the initial turnover at infinitely small time scales.

Based on this, I fit MM equations to the simulated data in order to test whether this would allow generating HD and affinity estimates that are independent from (or less affected by) duration and sampling completeness but reflect other aspects that are important to us.

In this markdown, I summarize the findings. The code for generating the HD and affinity estimates is from the accompanying R-script ("affinity.R"). This document only gives main results. I use the neutral simulations V2 and the four levels of sampling (25, 50, 75, 100). We have 6188 unique combinations of time series and parameter combinations for each of these subsets.

Outcome

Goodness of fit

One concern is that a large fraction of models does not converge on reliable estimates. To check this, I give a table with the number of positive estimates, no convergence, or even negative aspects.

Table 1: Fraction % of 6188 time series with positive or negative estimates for affinity and halving distance for Jaccard and Morisita-Horn

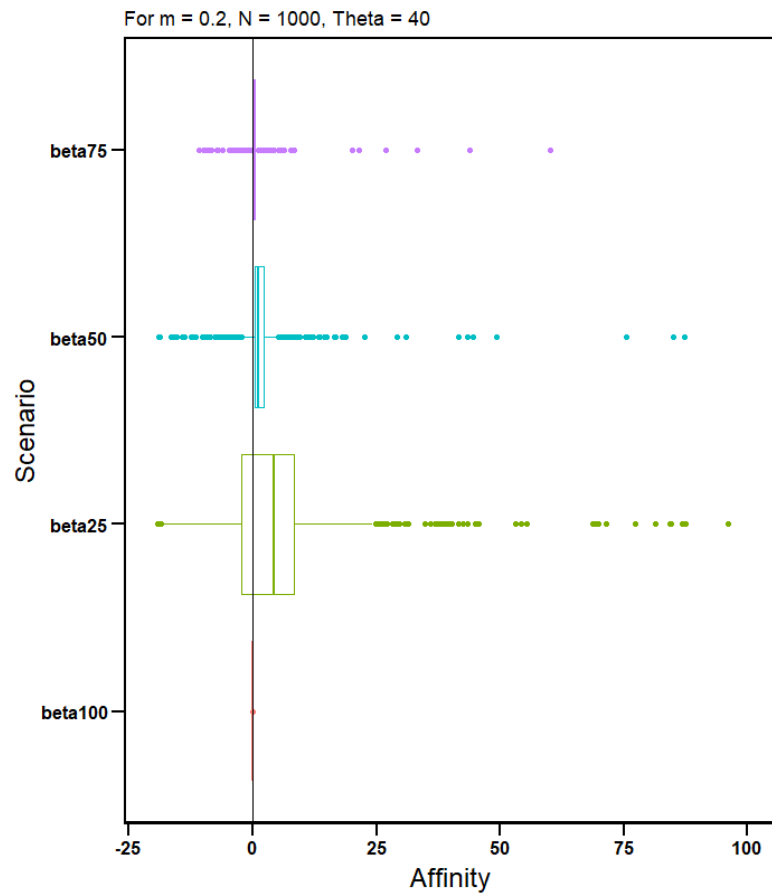|	Jaccard			Morisita		
scenario	pos_jac	neg_jac	na_jac	pos_mh	neg_mh	na_mh
beta100	58.14	0.00	41.86	50.11	0.00	49.89
beta25	67.34	32.26	0.40	66.77	32.82	0.40
beta50	76.49	23.25	0.26	77.68	22.03	0.29
beta75	84.97	14.75	0.27	89.27	10.57	0.16

Fig. 1: Estimates of affinity for the four sampling efforts (some extreme values < -25 or > 100 excluded).

First of all, we see a strange behaviour at 100% sampling, which leads to quite a substantial fraction of non-converging regressions. This fraction is negligible for the other sampling intensities, which instead have negative estimates. These negative estimates become less frequent when going from 25% to 75% sampling effort. At 75% sampling 85-90% of the regressions converge.

Plots

For the following plots, I use (log+1)-transformed estimates of affinity and HD. Therefore, the negative estimates and non-converging regressions are excluded.
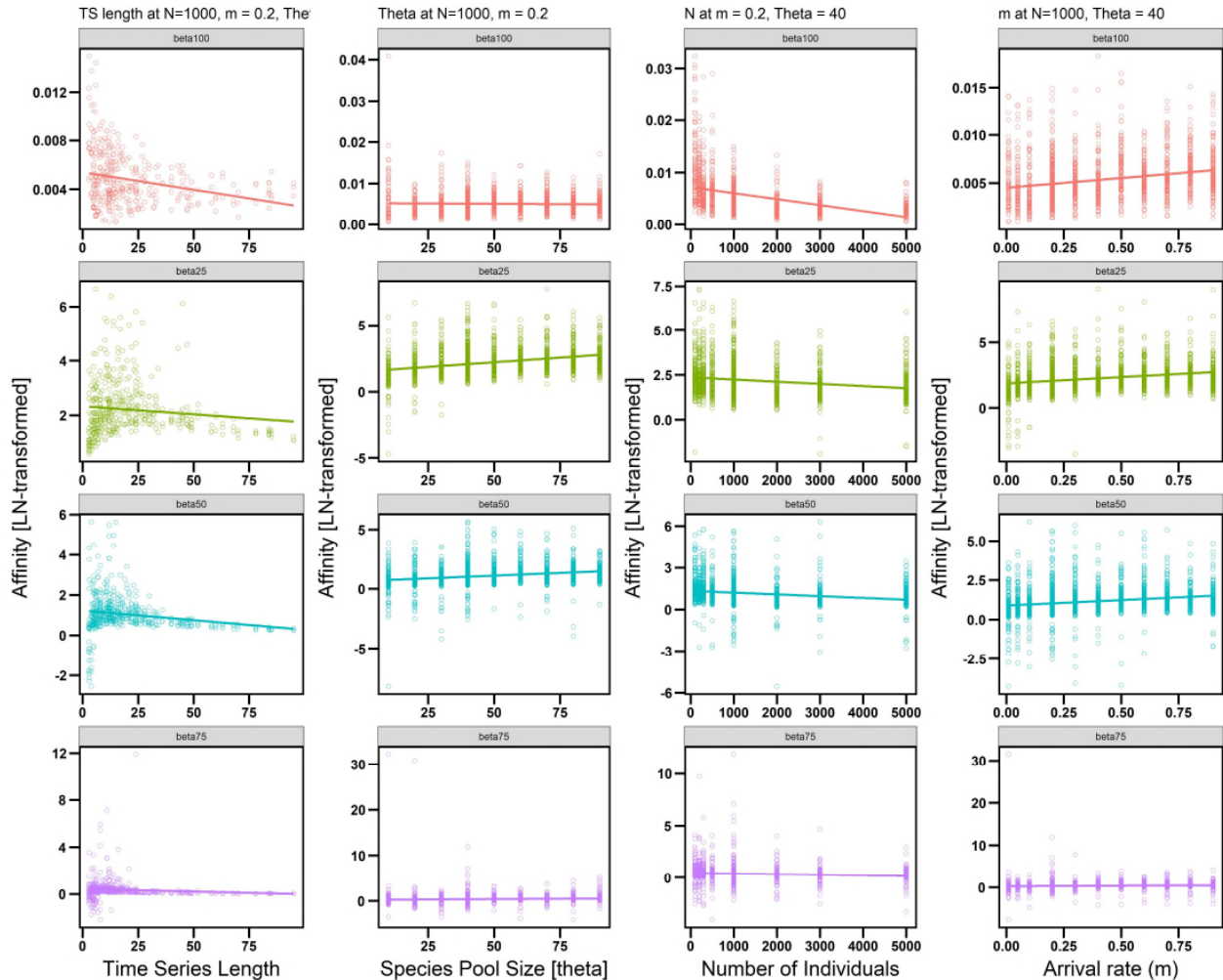

Fig. 2: Affinity on Jaccard basis in response to time series lengths, species pool size, number of individuals, and colonization rate.

Time series length has a stronger impact on affinity at low sampling interval, but at 75% there is almost no change, also the species pools size, N or m have little impact.
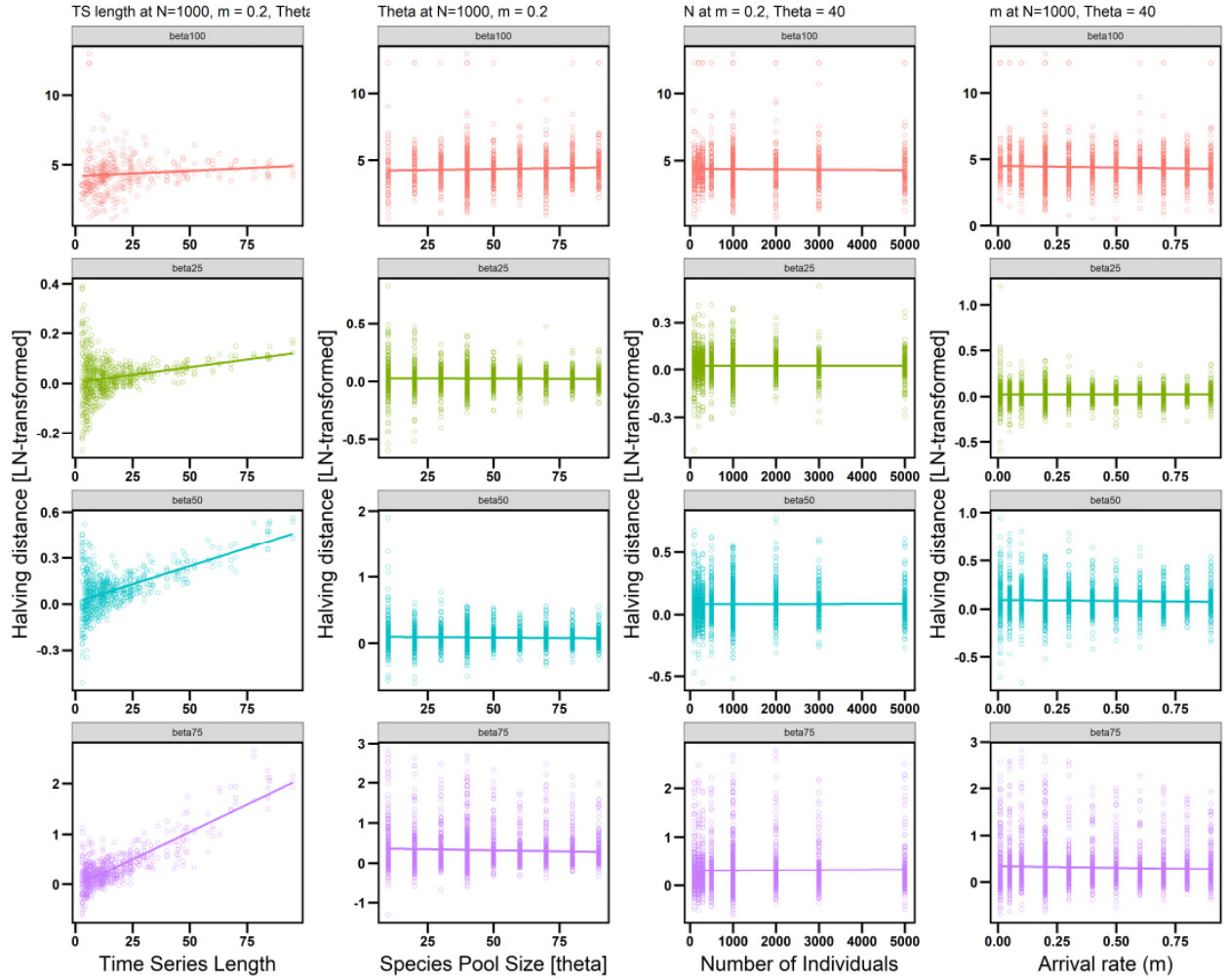
Fig. 3: Halving distance on Jaccard basis in response to time series lengths, species pool size, number of individuals, and colonization rate.

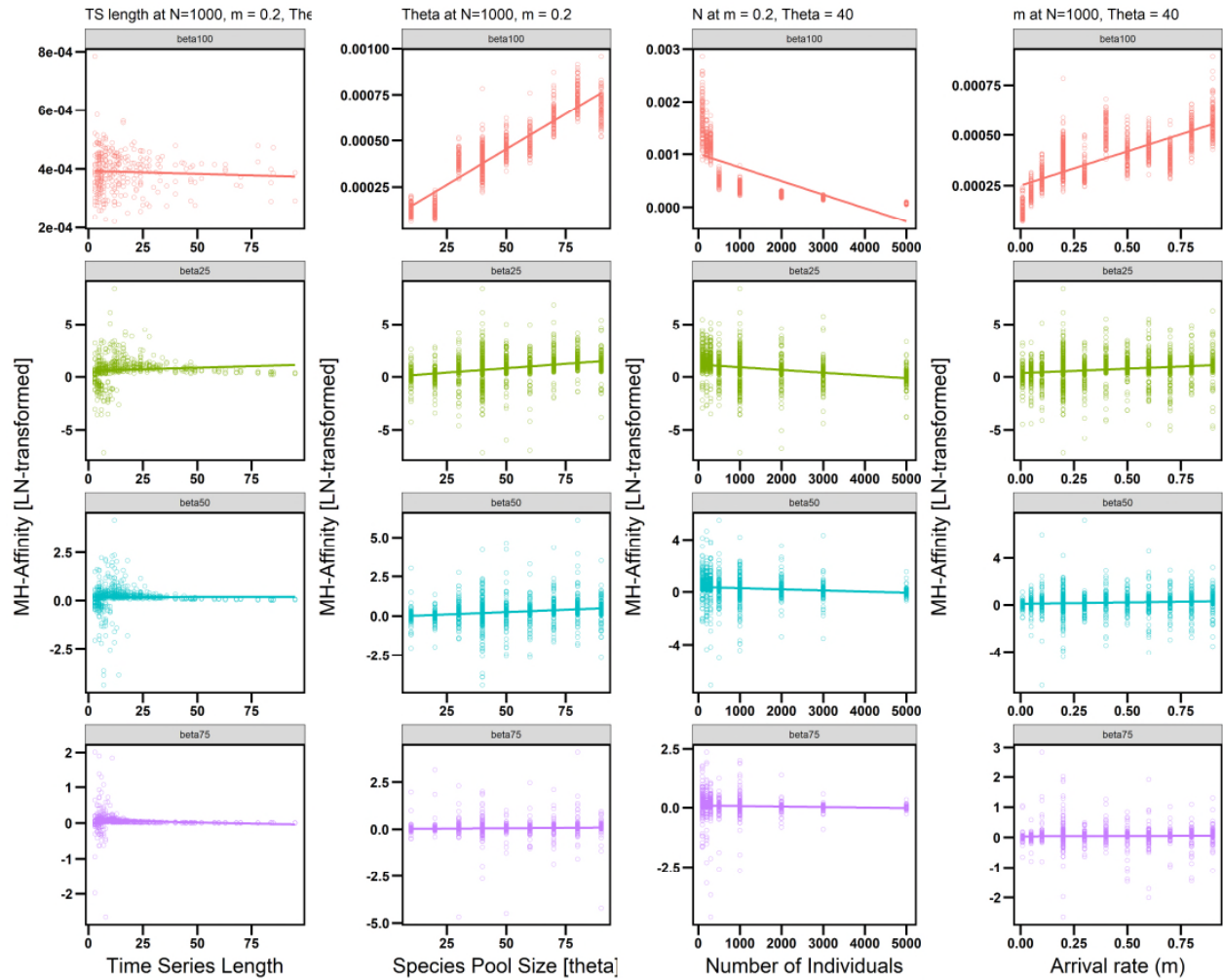HD is not affected by N, theta and m, but strongly by time series length.

Fig. 4: Affinity on Morisita Horn basis in response to time series lengths, species pool size, number of individuals, and colonization rate.

Time series length has almost no impact on MH-based affinity, but at 100% there are strong impacts of species pools size, N or m have little impact. These are minor at 75%.
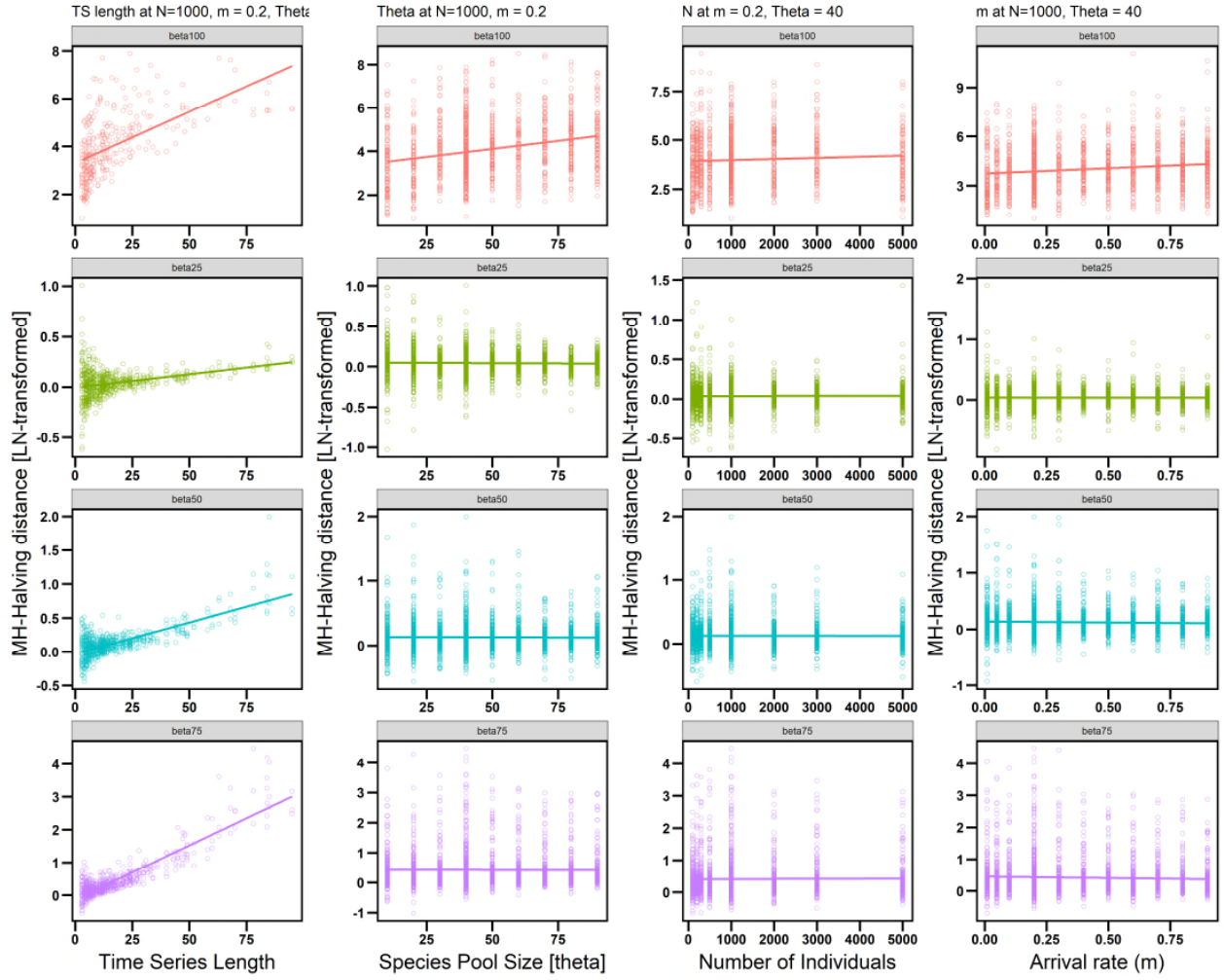
Fig. 5: Halving Distance on Morisita basis in response to time series lengths, species pool size, number of individuals, and colonization rate.

MH-based HD is very variable, but is not systematically affected by N, theta and m, but strongly by time series length.