

Katahdin power analysis

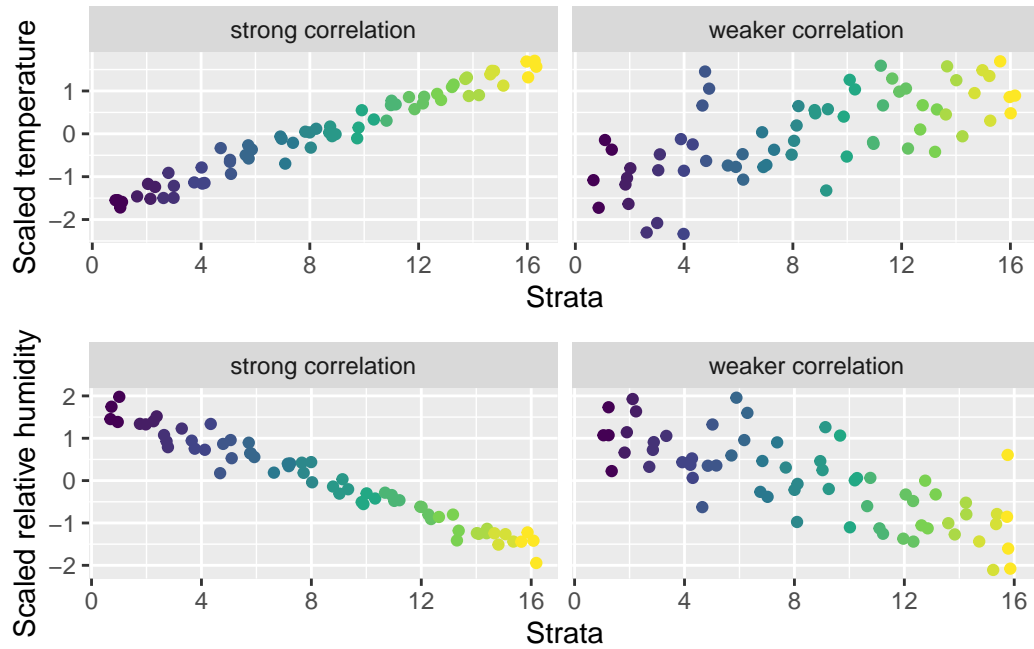
The purpose of this power analysis is to determine sufficient sample size on Katahdin. To do this, we will use as response variable the percent cover of hypothetical exemplar taxa. We will assume these hypothetical taxa have a climatic optimum where their percent cover peaks. We will deem sample size sufficient if that sample size has adequate power to detect the climatic optimum.

Relationship between strata and climate variables

We will consider two climate variables: temperature (**tmp**) and relative humidity (**rh**). There are 16 strata and we will investigate two different scenarios:

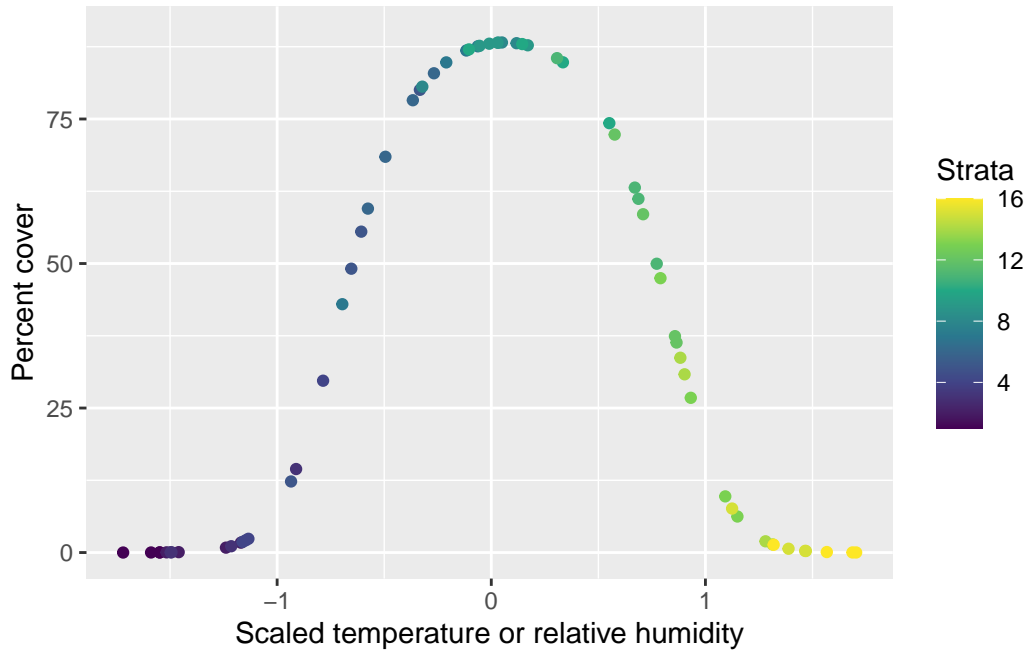
1. **tmp** and **rh** have strong correlation to strata
2. **tmp** and **rh** have weaker correlation to strata

The scenario of no correlation is unlikely given that strata were chosen to represent meaningful climatic variation.



Setting up possible scenarios for the relationships between focal taxa and climate

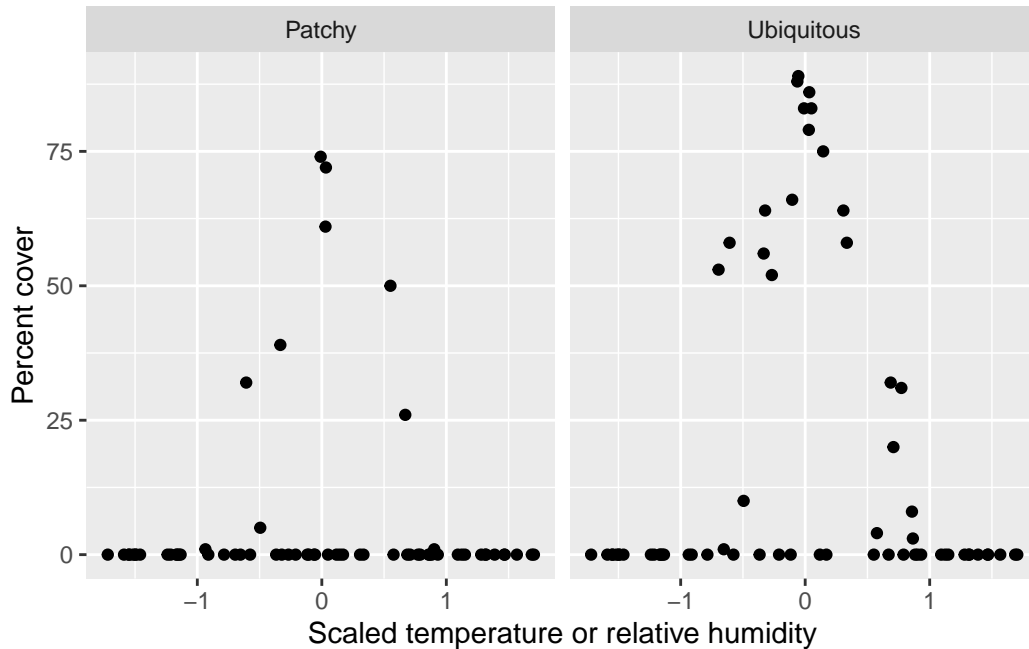
We will assume there is an optimal climate for the taxon where its percent cover peaks as displayed in the below figure which shows an idealized response of percent cover to climate (i.e. without random noise).



Different taxa will have different optimal climates, different maximum percent covers, and different tolerances for climatic conditions away from their optimal climate (i.e. narrow or wider peaks in percent cover).

The goal of our power analysis is to determine what sample size is necessary to accurately capture the climatic preferences of focal taxa under different representative scenarios of patchy versus ubiquitous spatial distributions and climate niches that are narrow versus wide.

We will model the patchiness of taxa via a zero inflated binomial distribution. The below figure shows the difference between a patchy spatial distribution and a ubiquitous spatial distribution.



What we can see is that the patchy taxon still peaks at the same climate optimum, but there are more sites where it was absent due to its patchy distribution. The ubiquitous taxon has fewer sites where it is absent, but there are still some gaps because we assume all taxa have some amount of patchiness.

Power analysis for relationship between focal taxa and climate

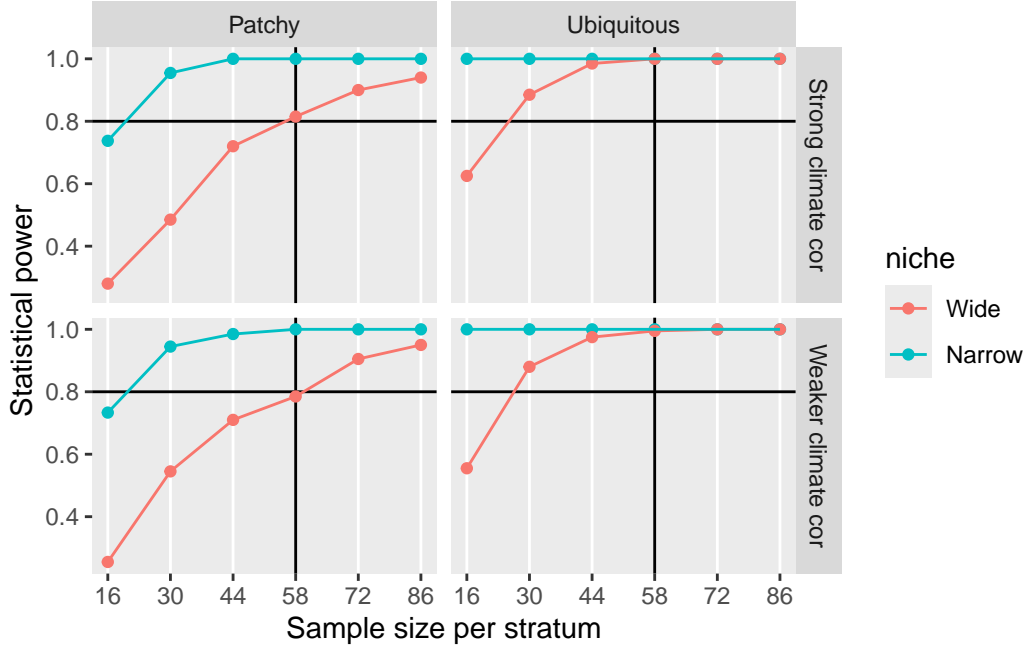
Now we investigate our statistical power to detect a response of the percent cover of a focal taxon to local climate variation, and how that power changes with number of sampling sites. We will investigate sample sizes ranging from 16 to 86. Sample sites are evenly spread across strata so that each stratum has nearly equal representation.

To recap, across those various sample sizes, we are investigating 8 scenarios coming from all unique combinations of the following conditions:

- *Variation in climate across strata:*
 1. temperature and relative humidity have strong correlation with strata
 2. temperature and relative humidity have weaker correlation with strata
- *Climate niche:*
 1. narrow
 2. wide
- *Spatial distribution:*

1. patchy
2. ubiquitous

The below figure shows how statistical power increases with the number of sample replicates per stratum under all combinations of those conditions.



The horizontal black line shows the 0.8 power cutoff. This is a commonly used cutoff for determining sufficient sample size. The vertical black line is our proposed sample size of 58. What we see is that 58 samples is close, if not exactly, the minimum number of points we need to reach a power of 0.8 across all combinations of conditions.

Interestingly enough, the strength of correlation between local climate and strata has effectively no impact on power. This is a very positive outcome because we do not know precisely what the correlation is between strata and climate *a priori*, but our power analysis indicates that this uncertainty is not detrimental to our study design.