 DOI: 10.1086/674525

# Appendix A from A. L. Hargreaves et al., "Are Species' Range Limits Simply Niche Limits Writ Large? A Review of Transplant Experiments beyond the Range"

(Am. Nat., vol. 183, no. 2, p. 157)

## Figures of Results Summarized in the Text of Hargreaves et al. (2014)

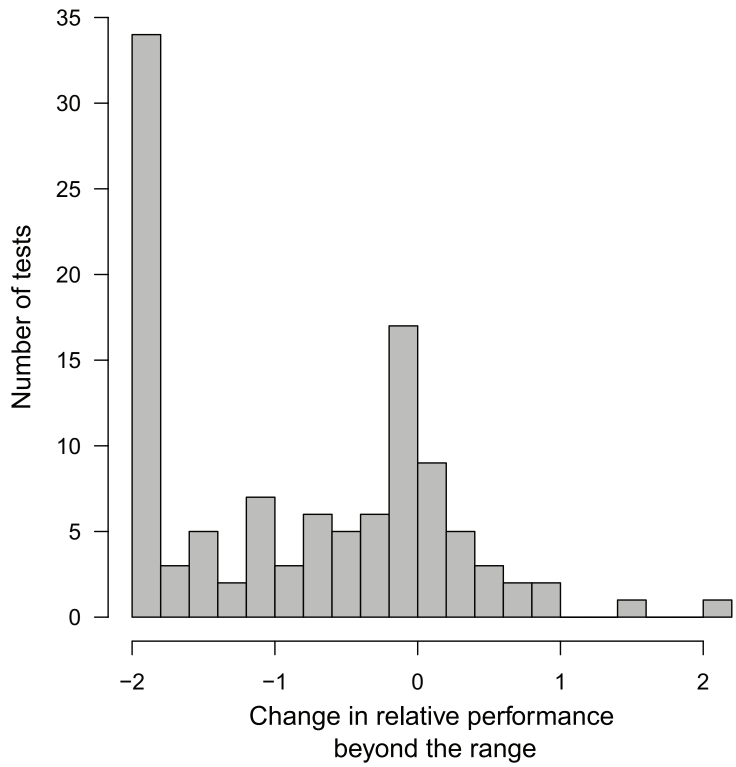


**Figure A1:** Distribution of the relative change in performance beyond the range limit, using the best comparison and best performance parameter for each of 111 tests. Each test is one taxon (species or subspecies) transplanted across one range limit. The change in relative performance is calculated as in table 1. Of 111 tests, 75 compared performance of source populations from the range interior (*i*) at transplant sites in the range interior (*I*) to sites beyond (*B*) the range (i.e., *i.BI*), 27 compared performance of interior sources between range-edge (*E*) and beyond-range sites (*i.BE*), and 9 compared performance of edge sources (*e*) between range-edge and beyond-range sites (*e.BE*). A negative number indicates performance declined beyond the range; −2 indicates zero fitness beyond the range. See text for summary statistics.

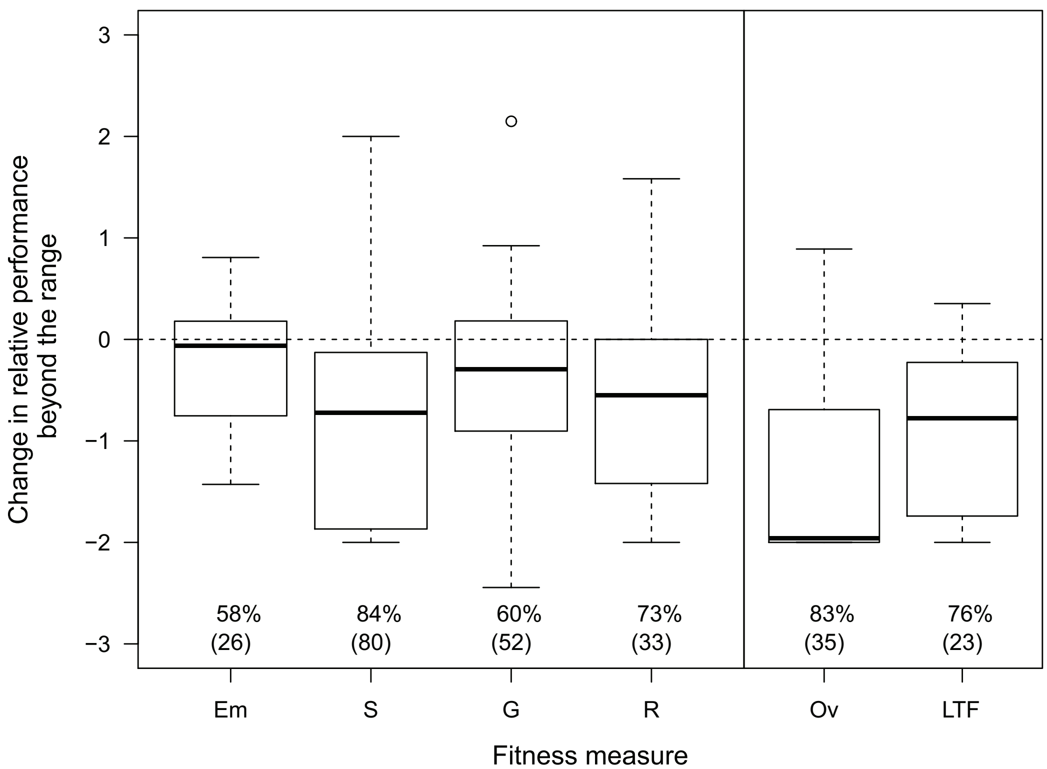


**Figure A2:** Performance declines beyond the range detected by each performance parameter (emergence [Em], survival [S], growth [G], reproduction [R]) and composite parameter (overall fitness [Ov]: product of two or more proportional fitness components, Em, S, R; lifetime fitness [LTF] as calculated by the authors or as the product of Em, S, and R when all three were available). Boxes indicate magnitude (median and quartiles, per fig. 3) of relative change in performance; negative numbers indicate performance was worse beyond the range than at the within-range (edge or interior) control site. Numbers below boxes indicate the percentage of tests in which performance declined beyond the range (number of tests for each parameter). Data include multiple parameters per test when available.

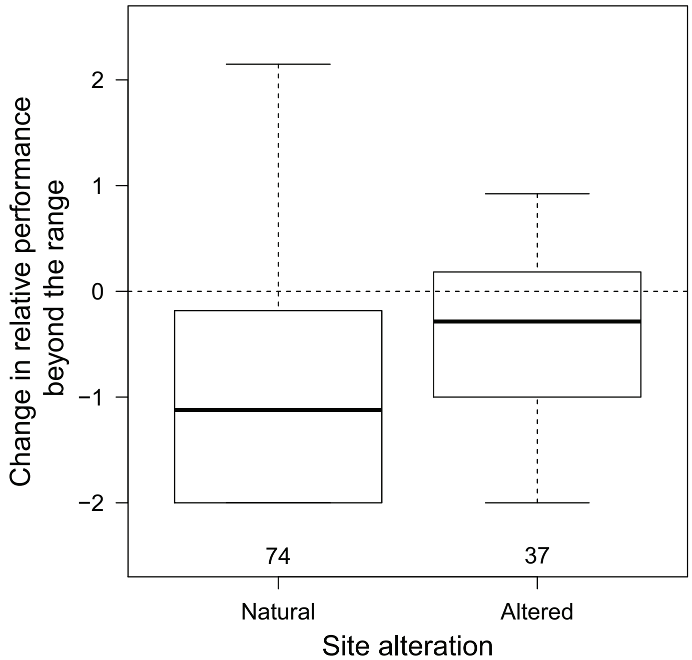


**Figure A3:** Change in relative performance beyond the range from transplant experiments conducted under natural conditions versus those in which one or more factors (e.g., competition, soil, watering) were standardized across transplant sites without control treatments. Means differ significantly (randomization $P \cong .0097$). Numbers below boxes are the tests in each category. The hatched line at zero indicates no difference in performance between individuals transplanted within the range (at interior or edge sites) and those transplanted beyond, and negative values mean fitness declined beyond the range. Boxplots show medians and quartiles as in figure 3.

 DOI: 10.1086/674525

# Appendix B from A. L. Hargreaves et al., "Are Species' Range Limits Simply Niche Limits Writ Large? A Review of Transplant Experiments beyond the Range"

(Am. Nat., vol. 183, no. 2, p. 157)

## Supplementary Analyses Testing for Experimental Bias in the Location and Replication of Beyond-Range Transplant Sites

### Comparing Geographic and Elevational Range Limits

Experimental biases could inflate the likelihood of finding stronger performance declines beyond elevational versus geographic range limits. Elevational studies may tend to place beyond-range sites climatically farther beyond the range because of the shorter physical distance needed to traverse elevational versus geographic climate gradients. If range limits occur across a continuous environmental gradient, then the distance between beyond-range transplants and within-range control sites may determine the magnitude of the performance decline detected. This, in turn, could increase the likelihood of determining that beyond-range sites were unsustainable and therefore of concluding that range limit (RL) is equal to niche limit (NL; RL = NL) instead of RL < NL (the likelihood of determining that RL > NL depends on zero fitness at edge sites and so should not be affected by such a bias in how far beyond sites are placed from the range limit).

To test for a bias in the climatic distance covered by elevational versus geographic studies, we determined whenever possible the distance (in kilometers for geographic ranges and meters above sea level [m asl] for elevational ranges) between (*a*) the range limit and the nearest beyond-range transplant site ($n = 99$ tests) and (*b*) the mean location of beyond-range sites and the mean location of the within-range sites they were compared to in the best comparison ($n = 109$ tests). Although converting between elevation and latitudinal changes in climate is difficult, a simplistic estimate is that mean air temperature decreases by 1°C for every ~100 m gain in elevation and every 100–145 km gain in latitude (Rohde 2012, as cited in Colwell et al. 2008). We therefore converted elevation differences to kilometers as 100 m asl = 100 km latitude (results did not differ when we used a more liberal conversion of 100 m asl = 145 km [not shown]). We then tested whether elevational and geographic studies differed in each measure using a randomized *t*-test (see "Comparing Performance Declines" in the main text for explanation of randomization tests). As expected, tests of elevational limits placed beyond-range sites climatically farther from both the range limit ([median, quartiles] elevation: 166, 64.5–160 km equivalents; geographic: 65.3, 40.0–66.0 km equivalents; *t*-test on $\log_{10}$ [distance], $P < .0001$) and within-range controls ([median, quartiles] elevation: 300, 217–528 km equivalents; geographic: 225, 40.0–380 km equivalents; *t*-test on distance, $P < .001$).

Next, we tested whether the distance between beyond-range and within-range sites used in the best comparison correlated negatively with the magnitude of beyond-range performance changes (i.e., the greater the distance between sites, the more performance declined beyond the range) and whether such a pattern might account for steeper fitness declines detected across elevational limits. We modeled the relative performance change beyond the range as a Gaussian response in a linear model, using a Pearson correlation to test the effect of distance across all tests and an analysis of covariance (ANCOVA) to contrast the effects on elevational and geographic range limits. Randomization tests were used to evaluate each of the potential effects in the models.

Distance between sites in the comparison was not correlated with the change in performance beyond the range when all studies were considered together ($r = 0.112$, df = 106, randomization $P \cong .126$). However, the ANCOVA revealed contrasting relations between elevational and geographic studies (distance × limit-type interaction: $F_{1,104} = 4.10$, $P \cong .0450$). Whereas distance between beyond-range and within-range sites in the best comparison was not related to how much performance declined across elevational limits ($r = 0.131$, df = 73, $P \cong .133$), it was positively correlated with the change in performance across geographic limits ($r = 0.573$, df = 31, $P \cong .0003$). In other words, studies that placed within and beyond sites farther apart tended to find smaller declines in performance beyond geographic range limits, the opposite of expectations. Even controlling for the effect of climatic distance using kilometer equivalents between sites in the best comparison, elevational studies detected significantly greater fitness declines beyond range limits (ANCOVA effect of limit type: $F_{1,104} = 13.2$, $P \cong .0003$). Thus, although the potential for bias exists (elevation studies place beyond-range sites further than geographic studies), there is no evidence that this bias accounts for the observed difference in performance decline between geographic and elevation range limits.

What might explain the counterintuitive positive correlation between performance change and climatic distance among geographic studies? We suspected that this relationship reflects methodological practices. Rather than greater climatic distances resulting in smaller performance declines, it seems more likely that studies conducted across shallow environmental gradients tend to both (*a*) place their sites farther apart to be sure to capture some degree of environmental variation and (*b*) detect smaller fitness declines. We tested this indirectly in two ways. First, we narrowed the analysis down to a comparison of high-elevation and high-latitude range limits, eliminating studies of eastern and western range limits. As climate does not change predictably across longitude the way it does across latitude and elevation, a given distance likely produces very different climatic changes across different longitudinal gradients, and thus we would expect the methodological bias of farther distances across shallower gradients to be strongest for longitudinal studies. Indeed, once longitudinal studies were removed, an ANCOVA found no relationship between distance from within-range to beyond-range sites and fitness declines beyond the range ($F_{1,56} = 0.197$, $P \cong .67$), and this did not differ between high-elevation and polar limits (distance × limit-type interaction: $F_{1,56} = 0.254$, $P \cong .620$). Even after controlling for the effect of distance, fitness declines were significantly greater across high-elevation limits than polar limits (effect of limit type: $F_{1,56} = 18.79$, $P \cong .0001$).

Second, we used 15 tests that had more than one beyond-range (*B*) transplant site and tested whether performance differed more between within-range controls and the nearest versus the farthest *B* site using a paired *t*-test (see "Comparing Performance Declines" in the main text). Performance declined more to the farthest *B* site than to the closest for 11 of 15 tests (9 of 11 elevational tests and 2 of 4 geographic tests), and as expected, the magnitude of the performance comparisons was significant ($t_{14} = 2.49$, $P = .0259$). This affirms that range limits often overlie gradients of continuously declining habitat quality, such that greater distances along a given gradient result in greater fitness declines. However, it also affirms that not all range limits overlie such gradients, highlighting both the potential importance of gradients other than habitat quality (fig. 1) and the variability of habitat-quality gradients in nature.

Finally, using a *t*-test, we assessed whether RL = NL tests were associated with greater distances between *B* and within-range control sites than RL < NL studies. Although RL = NL studies were associated with higher distances between sites (median 208 km equivalents vs. 77 km equivalents for RL < NL studies), this difference was not significant ($t_{24} = -1.64$, $P \cong .113$, $n = 18$ tests). Given the lack of simple relation between distance and relative performance changes (above), we feel it is safe to assume that the differences detected between elevational and geographic limits are not due to simple methodological differences between studies but rather reflect steeper ecological gradients across elevation versus latitude.

## Replication of Beyond-Range Transplant Sites

Whereas suitable habitat within the range can be identified by the presence of the study species, identifying suitable habitat beyond it may be difficult unless the species has very specific and known habitat requirements. With a single beyond-range site, the risk of choosing an inappropriate site and mistakenly concluding that conditions beyond the range are unsuitable may be high. To test for such a bias, we compared performance declines beyond the range between tests that replicated beyond-range sites ($n = 46$) and those with only a single beyond-range site ($n = 65$). We used a 2 × 2 $\chi^2$ test to gauge whether the frequency of performance declines and randomization tests to gauge whether the magnitude of declines differed. The frequency of beyond-range performance declines did not differ between tests with only 1 beyond-range site (70.7%) and those with >1 (80.4%, $\chi^2_1 = 1.33$, $P = .29$). Contrary to expectations, performance declined more for tests with >1 *B* site (mean ± SD of relative performance change $-1.035 \pm 1.006$) than those with a single *B* site ($-0.677 \pm 0.955$), though this difference is not quite significant (randomization $P \cong .059$). Thus, although the methodological arguments for including >1 beyond-range transplant site are sound, there is no evidence that the meta-analysis results are biased by a common failure to do so.

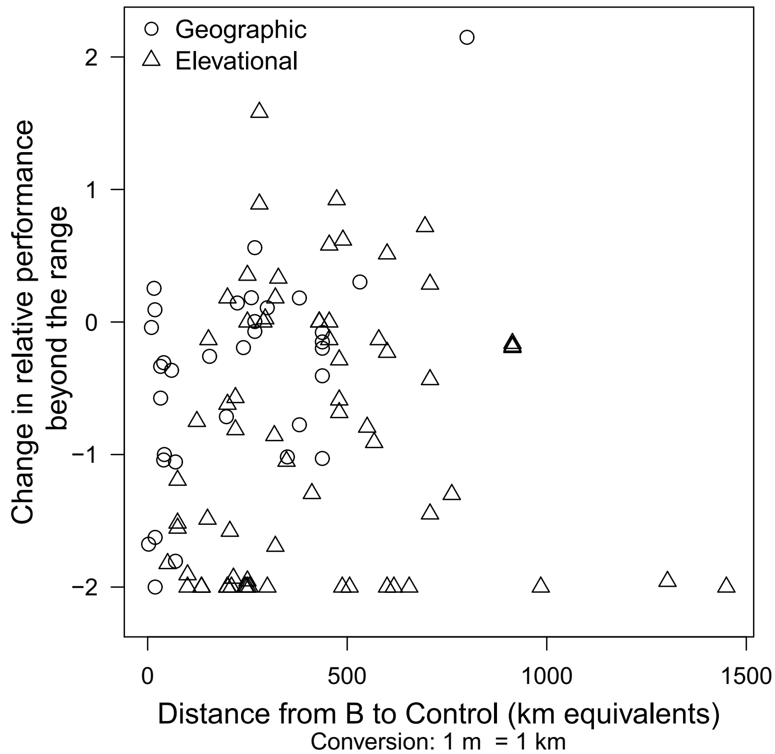


**Figure B1:** Distance from the beyond-range planting site to the within-range planting site used in the best comparison is not correlated with the change in performance beyond the range detected by that comparison ($r = +0.148$, randomization $P \cong .062$). Relative performance differences consider the best measure of performance and the best site comparison (see text) for each test. Changes in elevation for elevational range limits were converted to kilometer equivalents as 1 m asl = 1 km latitude.