modern-day female authors who situate their plots in the colonial period. In 2013, she published, *Llámenme «el mexicano»: Los almanaques y otras obras de Carlos de Sigüenza y Góngora* (Peter Lang). She has also published short stories. During the summer of 2013, she spent time at Seoul's National University, and in summer 2014, in Kyungpook National University, both in South Korea.

Statistical Analysis of Evergreen Invaders

Michael Lloyd, Ph.D.
Professor of Mathematics

Jonathan Eagle, B.S.
Biology

Abstract

The reproductive status, height, and distribution of seven types of invasive evergreens were analyzed. In Arkansas, about 23–26% of the flora consists of non-native species (Arkansas Vascular Flora Committee 2006). Some of the most invasive plants in the southeastern United States are woody ornamentals like the ones studied in this paper. This was a collaborative effort with Dr. Brett Serviss.

Introduction

The number of cases was 5765 and the variables were area (Arkadelphia, Hot Springs), site (1–46), species (*Elaeagnus pungens*, *Ilex cornuta*, *Ligustrum japonicum*, *Ligustrum lucidum*, *Mahonia bealei*, *Nandina domestica*, *Photinia serratifolia*), reproductive (yes, no), and height in centimeters. It was assumed that it was unnecessary to consider the site variable in any analysis. When a genus like *Elaeagnus*, *Ilex*, *Mahonia*, *Nandina*, or *Photinia* is mentioned in this paper, then it is referring to precisely one of the species listed above.

Figure 1: Non-reproductive

An evergreen is considered reproductive if and only if berries are present. Figure 1 shows an example of *Nandina domestica* that is not reproductive; Figure 2 shows an example of the same species that is reproductive.

Figure 2: Reproductive

Height versus species was graphed in Figure 3 in order to visualize the center, spread, and shape of the distribution of heights. Because the height distributions for most of the species have outliers or are skewed right, nonparametric procedures will be favored in subsequent analyses.

Figure 3: Height versus species

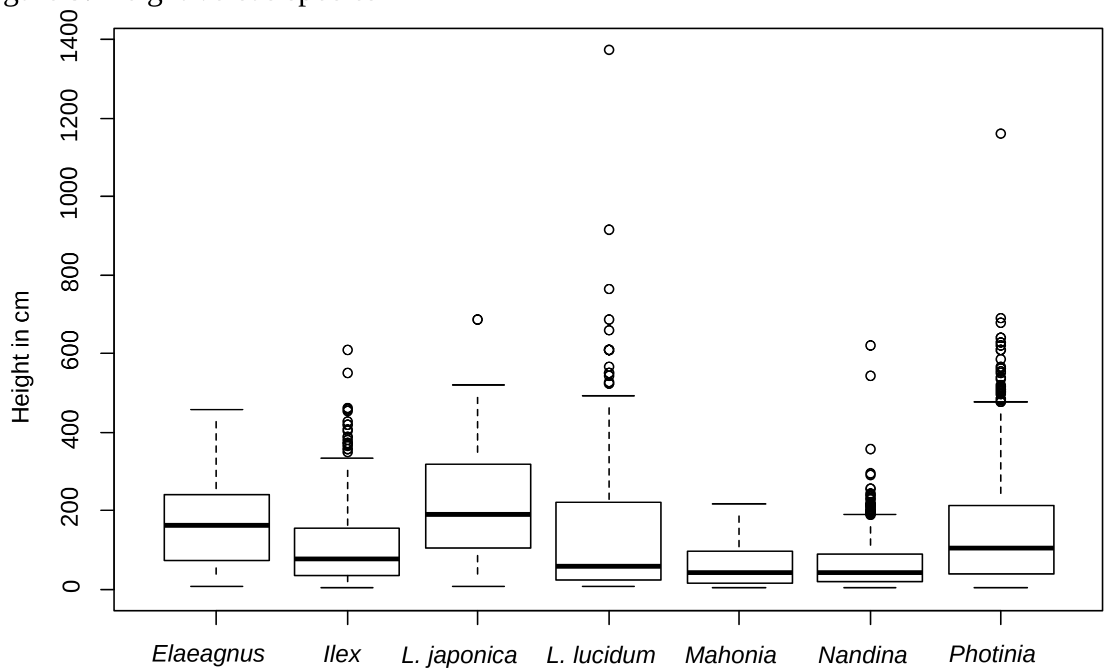


Table 1 shows the summary statistics for each species. Because most of the heights appeared to be non-normally distributed, ordinal measures for the center and spread were used. The reproductive rate is defined as the number of reproductive plants divided by sample size.

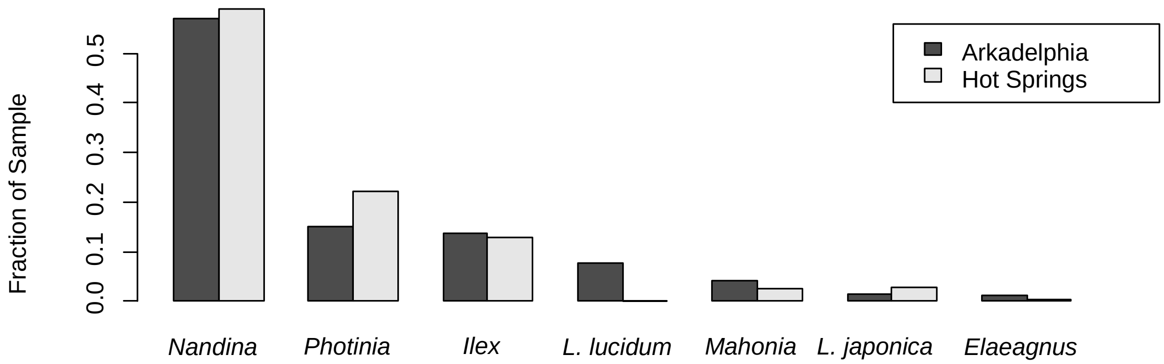
Table 1: Summary statistics

Species	Sample size	Median height (cm)	Interquartile range (cm)	Reproductive rate
Elaeagnus	49	163	165	0.18
Ilex	770	75	123	0.07
L. japonica	119	188	216	0.13
L. lucidum	264	59	198	0.05
Mahonia	197	42	83	0.31
Nandina	3327	43	68	0.38
Photinia	1039	102	176	0.07

Sample size dependence on area

We will first investigate if the distributions of samples for the species depend on the area. It is particularly important that the samples were chosen to represent the fraction of species within each area. Figure 4 shows that the distributions appear approximately the same, except for *L. Lucidum*, which was almost absent in the Hot Springs area.

Figure 4: Fraction of sample versus species by area



However, a chi-squared test of homogeneity yielded $\chi^2 = 249$, $df = 6$, $pvalue \approx 0$. Hence, there is strong evidence that the distributions of species were not the same for the Arkadelphia and Hot Springs areas. The standardized residuals are shown in Table 2. For *Nandina* and *Ilex*, these were relatively small (less than 2). This indicates that the fraction of these species were not apparently different than expected if, in fact, the distribution of species were the same for these two areas. However, the proportion of the other species in the sample are different than expected.

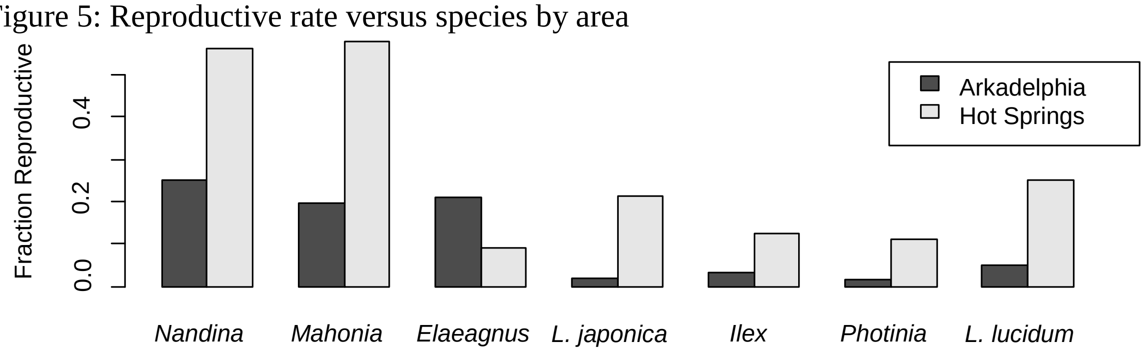
Table 2: Comparing species distribution between areas

Species	Arkadelphia sample size	Hot Springs sample size	Arkadelphia fraction	Hot Springs fraction	Absolute standardized residual	Apparently different
Nandina	1912	1415	0.569	0.589	1.5	no
Photinia	507	532	0.151	0.222	6.9	yes
Ilex	457	313	0.136	0.130	0.6	no
L. lucidum	260	4	0.077	0.002	13.6	yes
Mahonia	138	59	0.041	0.025	3.4	yes
L. japonica	49	70	0.015	0.029	3.8	yes
Elaeagnus	38	11	0.011	0.005	2.7	yes
Total	3361	2404	1.000	1.000	—	—

Reproductive rate dependence on area

It appears from Figure 5 that, except for *Elaeagnus*, every species is more likely to be reproductive in Hot Springs than it is in Arkadelphia.

Figure 5: Reproductive rate versus species by area



A 2-proportion test with continuity correction was done for each species. The p-values may be inaccurate for *Elaeagnus* and *L. lucidum* because they had small samples sizes of 11 and 4 from Hot Springs, respectively. The results in Table 3 suggest that the reproductive rate was greater near Hot Springs than near Arkadelphia for all of the species except *Elaeagnus* and *L. lucidum*. Larger sample sizes may have provided significant p-values for these species.

Table 3: Comparing reproductive rates between areas

Species	**Arkadelphia sample size**	**Hot Springs sample size**	**Arkadelphia reproductive rate**	**Hot Springs reproductive rate**	**2-sided p-value**	**Significantly higher rate near Hot Springs**
Nandina	1912	1415	0.25	0.56	≈ 0	yes
Mahonia	138	59	0.20	0.58	≈ 0	yes
Elaeagnus	38	11	0.21	0.09	0.65	no
L. japonica	49	70	0.02	0.21	0.005	yes
Ilex	457	313	0.03	0.12	≈ 0	yes
Photinia	507	532	0.02	0.11	≈ 0	yes
L. lucidum	260	4	0.05	0.25	0.52	no

Height dependence on area

Figures 6–12 show boxplots for the height dependence on area for each species. Except for *Ilex*, the species tended to be taller in Hot Springs (H) than in Arkadelphia (A).

Figure 6: *Nandina* area vs. height

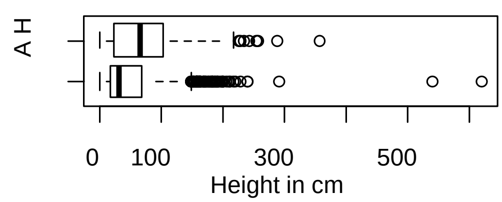


Figure 7: *Mahonia* area vs. height

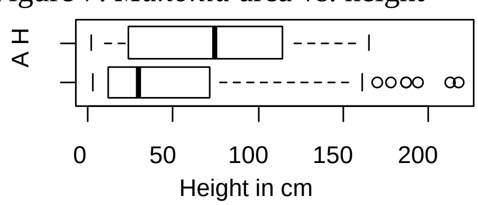


Figure 8: *Elaeagnus* area vs. height

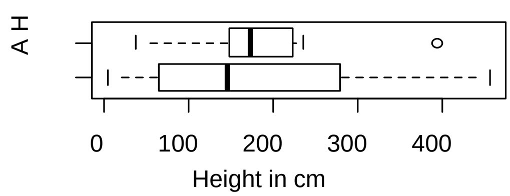


Figure 9: *L. Lucidum* area vs. height

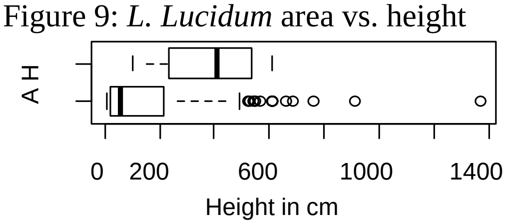


Figure 10: *Ilex* area vs. height

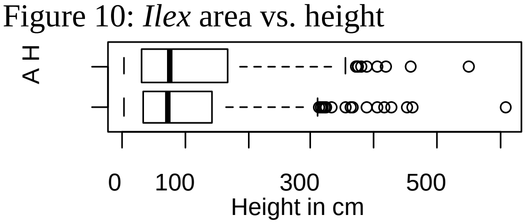


Figure 11: *L. Japonica* area vs. height

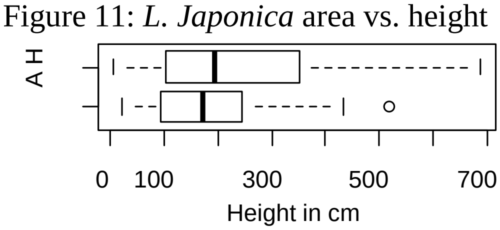
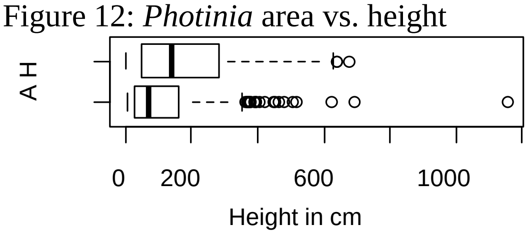


Figure 12: *Photinia* area vs. height



See Table 4 for the results of a Wilcoxon Rank-Sum Test applied to the each of the species. These tests indicated that four of the seven species were systematically taller near Hot Springs than near Arkadelphia.

Table 4: Comparing heights between areas

Species	Sample size in H	Sample size in A	Median height in H	Median height in A	H−A Location (cm)	Wilcoxon Rank-Sum W statistic	2-sided p-value	Significantly different
Nandina	1415	1912	66	33	19	1,032,492	≈ 0	yes
Mahonia	59	138	75	29.5	26	3018	0.004	yes
Elaeagnus	11	38	174	146	15	189.5	0.65	no
L. lucidum	4	260	411.5	58	303	187	0.03	yes
Ilex	313	457	76	74	4	69,079	0.42	no
L. japonica	70	49	196.5	174	26	1527	0.31	no
Photinia	532	507	137.5	69	47	98,155	≈ 0	yes

Height dependence on reproductive rate

Figures 13–19 show boxplots for the reproductive rate dependence on height. In every case, the reproductive plants tended to be taller. (N is non-reproductive, and R is reproductive.)

It is possible that the populations in Hot Springs are more mature than the ones in Arkadelphia. The median height for reproductive *L. lucidum* was 610 cm and its interquartile range was 0 cm.

Figure 13: *Nandina* reproductive vs. height

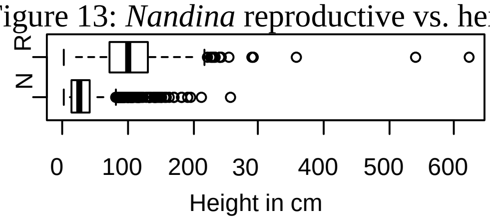


Figure 14: *Mahonia* reproductive vs. height

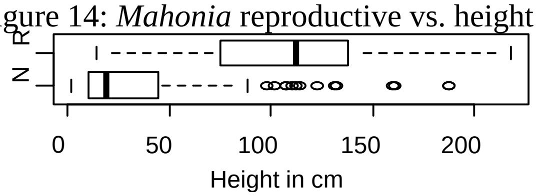


Figure 15: *Elaeagnus* reproductive vs. height

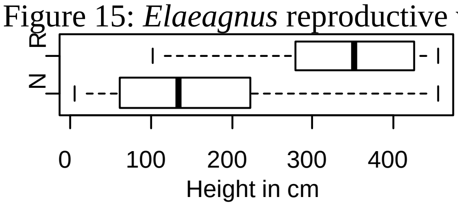


Figure 16: *L. Lucidum* reproductive vs. height

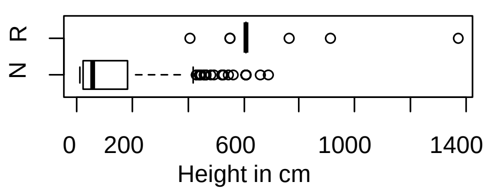


Figure 17: *Ilex* reproductive vs. height

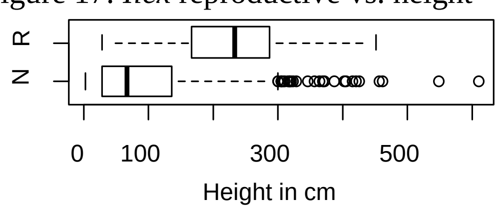


Figure 18: *L. Japonica* reproductive vs. height

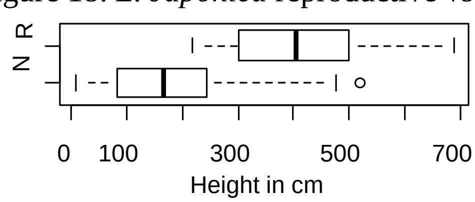
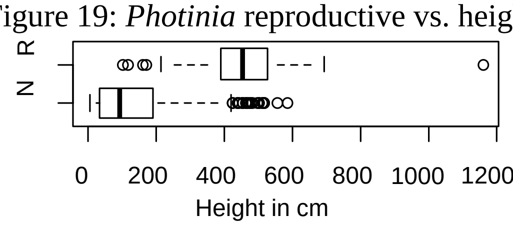


Figure 19: *Photinia* reproductive vs. height



Height appears to have a stronger influence on reproductive rate than area. Table 5 shows the results of Wilcoxon Rank-Sum Tests. It was found that the reproductive plants were systematically taller than the non-reproductive plants for all of the species.

Table 5: Comparing heights for the two reproductive states

Species	**Sample size of R**	**Sample size of N**	**Median height of R (cm)**	**Median height of N (cm)**	**Height difference (R−N) location (cm)**	**Wilcoxon Rank-Sum W statistic**	**2-sided p-value**	**R significantly taller than N**
Nandina	1275	2052	99	25	69	180,002	≈ 0	yes
Mahonia	61	136	112	18.5	77	796	≈ 0	yes
Elaeagnus	9	40	352	134	192	55	0.001	yes
L. lucidum	14	250	610	55.5	548	102	≈ 0	yes
Ilex	54	716	235	69	149	5338	≈ 0	yes
L. japonica	16	103	403.5	162	230	190	≈ 0	yes
Photinia	68	971	449.5	89	345	3148	≈ 0	yes

Logistic models for the reproductive rates

Since reproductive rates have been shown to usually depend on height and area, we will attempt to find a logistic model for predicting the reproductive state for each of the species using these as the explanatory variables. To check that such a model is appropriate, see Figures 20–26 for empirical logistic plots for each of the species. The short *Nandina* that appear to not fit the linear logit model will be addressed later. The lines are from the logistic fit for height only; they are not regression lines.

Figure 20: *Nandina* logit plot

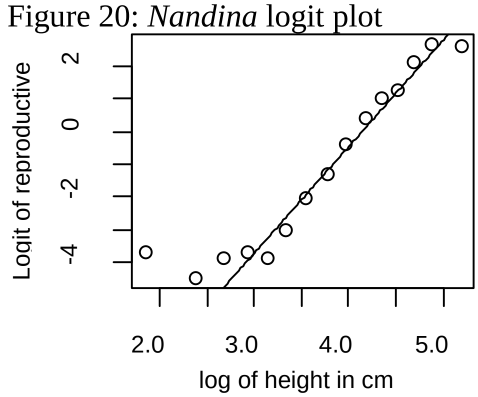


Figure 21: *Mahonia* logit plot

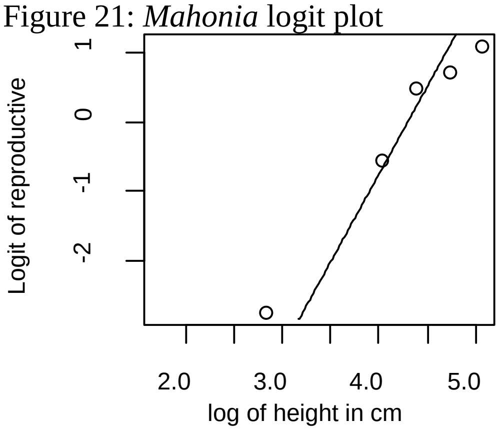


Figure 22: *Elaeagnus* logit plot

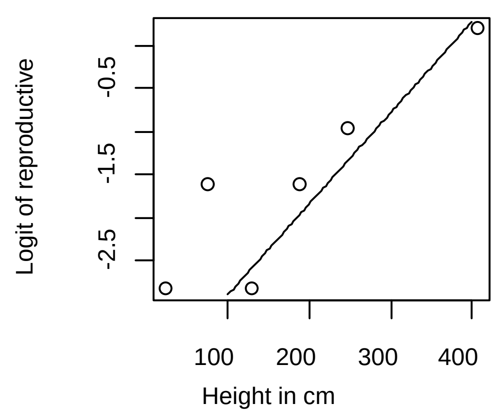


Figure 23: *L. lucidum* logit plot

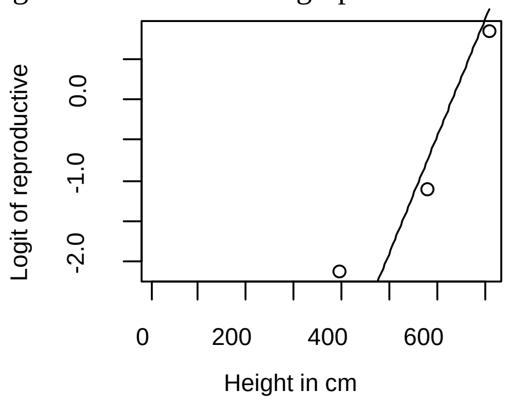


Figure 24: *Ilex* logit plot

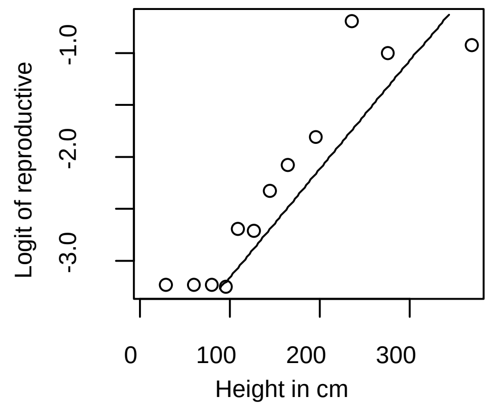


Figure 25: *L. Japonica* logit plot

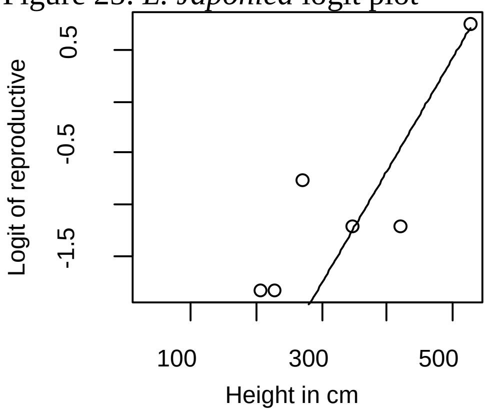
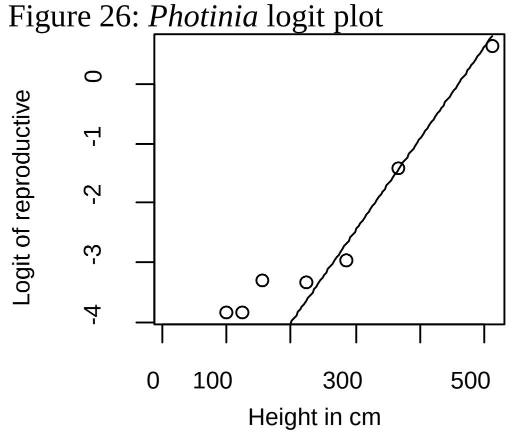


Figure 26: *Photinia* logit plot



Only the simple transformation $\log = \log_e$ was found to improve the fit for some of the logit plots. See Table 6 for a summary of the height transformations, AIC criteria, prediction success, and p-values for both variables.

Table 6: Variables in Logistic Regression

Species	**Sample size**	**Height transformation**	**AIC (height)**	**AIC (height & area)**	**P-value (height)**	**P-value (area)**
Nandina	3327	log	2209	2089	≈ 0	≈ 0
Mahonia	197	log	141	128	≈ 0	≈ 0
Elaeagnus	49	none	38	40	0.002	—
L. lucidum	264	none	49	50	≈ 0	—
Ilex	770	none	320	299	≈ 0	≈ 0
L. japonica	119	none	67	62	≈ 0	0.03
Photinia	1039	none	245	241	≈ 0	0.02

The Akaike Information Criteria (AIC) was used to decide if both the height and area, or only the height, should be included in the logistic model for predicting the reproductive rate. It was decided that only height should be used in the *Elaeagnus* and *L. lucidum* models. In Table 3, it was seen that area was not found to significantly affect the reproductive status for these species when a 2-proportion test was performed earlier. All of the other coefficients were significant.

Table 7 shows the coefficients in the model and the odds ratios. The area value is 1 for Hot Springs and 0 for Arkadelphia. The area terms were part of the models for *Nandina*, *Mahonia*, *Ilex*, *L. japonica*, and *Photinia*. This means that the probability of these species being reproductive were higher in Hot Springs than Arkadelphia in a way that could not be adequately explained by their heights alone. For example, for a *Nandina* of fixed height, its odds of being reproductive was 3.39 times higher near Hot Springs than if it were near Arkadelphia.

Table 7: Coefficients & Odds Ratios

Species	**Height transformation**	**Intercept**	**log(height) or height coefficient**	**Area coefficient (H=1)**	**Odds ratio (height in m)**	**Odds ratio (area)**
Nandina	log	-13.1	3.06	1.22	—	3.37
Mahonia	log	-11.3	2.42	1.75	—	5.76
Elaeagnus	none	-4.04	0.0109	—	2.98	—
L. lucidum	none	-9.20	0.0149	—	4.30	—
Ilex	none	-5.18	0.0193	1.57	2.98	4.79
L. japonica	none	-6.86	0.0107	2.43	2.92	11.3
Photinia	none	-7.55	0.0145	1.07	4.26	2.92

The odds ratios are only displayed when no transformation was applied to the height. Also, the units were changed to meters for the odds ratio because an additional 1 cm in height had little effect on the odds ratio. For example, if the height of *L. lucidum* were increased by 1 m, then the odds of this plant being reproductive would increase by a factor of 4.30 times.

Measuring model predictive accuracy

The measure we used for the predictive accuracy is called the c-statistic or c-index (Page 574, STAT2 2003). This considers all possible pairings of cases which were reproductive and those which were not. A concordance means that the model predicted a higher chance of being reproductive for the plants that actually were reproductive than the ones that were not. As you can see in Table 8, the number of possible comparisons using this method can be large. Our predictive models were considered to be successful if their c-statistics were at least 90%.

Table 8: Predictive accuracy of models

Species	**Sample size**	**Number not reproductive**	**Number reproductive**	**Number of comparisons**	**c-statistic**	**Model satisfactory**
Nandina	3327	2052	1275	2,616,300	93%	yes
Mahonia	197	136	61	8296	90%	yes
Elaeagnus	49	40	9	360	84%	no
L. lucidum	264	250	14	3500	96%	yes
Ilex	770	716	54	38,664	86%	no
L. japonica	119	103	16	1648	88%	no
Photinia	1039	971	68	66,028	95%	yes

The relatively large number of tall, non-reproductive outliers for *Ilex* as seen in Figure 17 contributed to it having the worst predictive rate. Too many cases for tall Ilex would have to have been removed to significantly improve its logistic model.

Figures 27–33 show probability of being reproductive plotted against the height for each species. For the models that included area as an explanatory variable, two curves are plotted with the following symbols for the empirical probability of being reproductive in the two areas:

□ = Hot Springs, △= Arkadelphia

If the model did not include area as an explanatory variable, then only ○ was used for a plot symbol.

Figure 27: *Nandina* reproductive probability

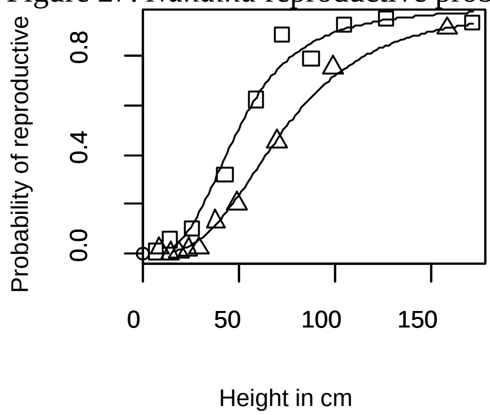


Figure 28: *Mahonia* reproductive probability

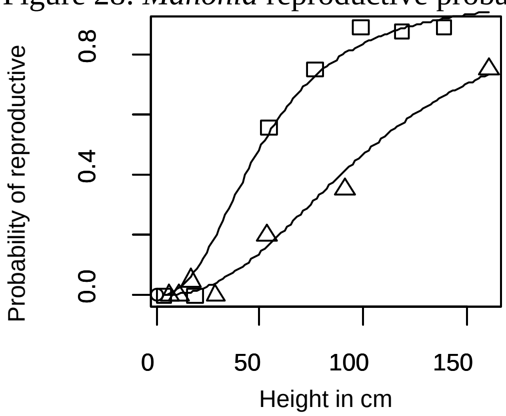


Figure 29: *Elaeagnus* reproductive probability

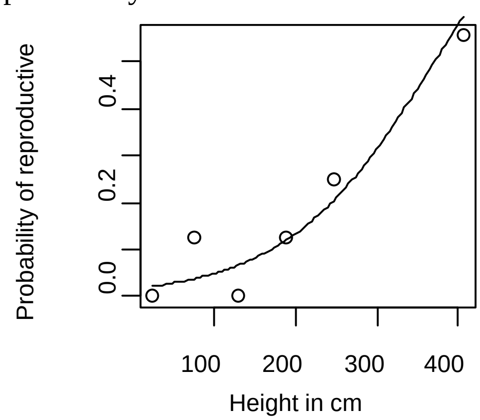


Figure 30: *L. lucidum* reproductive probability

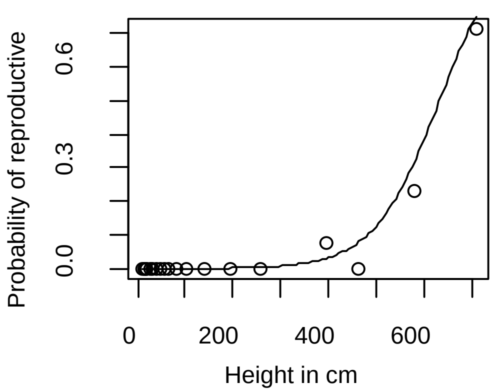


Figure 31: *Ilex* reproductive probability

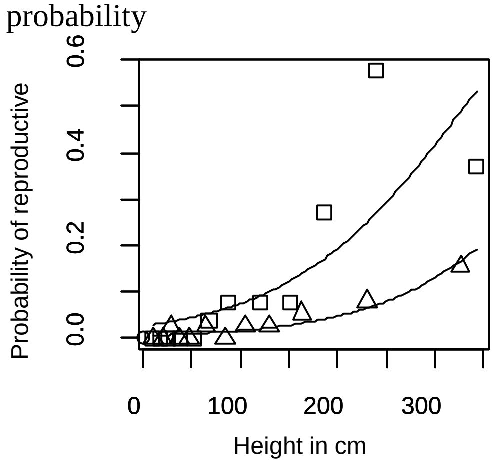


Figure 32: *L. japonica* reproductive probability

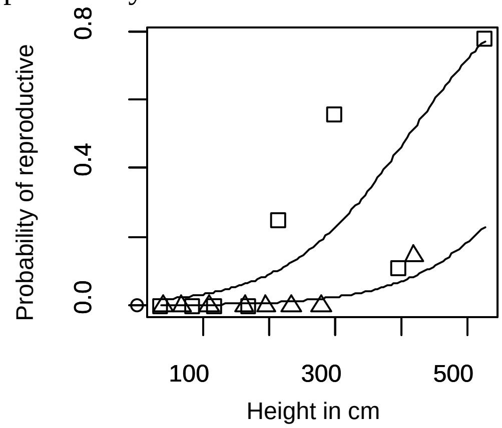
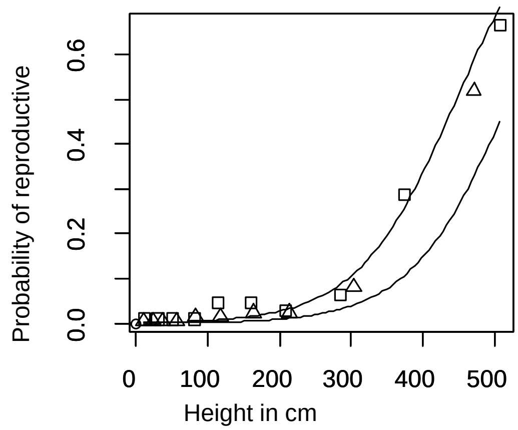


Figure 33: *Photinia* reproductive probability



Short *Nandina*

Refer to Figure 20 to see that there are shorter *Nandina* that do not appear to fall near the logit line. We will designate as short *Nandina* those that are shorter than 15 cm. See Table 9 for a summary of these plants.

Table 9: Summary of short *Nandina*

Sample size	Sample near Arkadelphia	Sample near Hot Springs	Median height near Arkadelphia (cm)	Median height near Hot Springs
529	303	226	10	8

The short *Nandina* that did not appear to fit the logistic model well were 16% of the original sample for this species. A logistic model constructed using the original 3327 *Nandina* was used to determine the reproductive probability of the short *Nandina* near Arkadelphia and those near Hot Springs. (The median heights for the short *Nandina* were used.) See Table 10 for the results of two exact Binomial Tests: There was not a significant difference between the observed and predicted probabilities of a short *Nandina* being reproductive for both the Arkadelphia and Hot Springs areas. Therefore, the *Nandina* data set should not be split into plants that were at least 15 cm tall and those that were less than 15 cm tall when constructing a logistic model.

Table 10: Binomial Tests comparing observed and predicted probability of being reproductive

Area	Sample size	Observed probability reproductive	Predicted probability reproductive	P-value
Arkadelphia	303	0.017	0.02	0.54
Hot Springs	226	0.027	0.02	0.47

Conclusion

The probability of being reproductive for all seven species were found to depend on height. The region (Hot Springs or Arkadelphia) significantly affected this probability for five of the species. All logistic models had a predictive accuracy of at least 84%; four of the logistic models had an accuracy of at least 90%.

Credits

We appreciate the efforts of Dr. Brett Serviss, who oversaw the project which was the source of the data used in this paper. Also, we appreciate the Ellis College Planning and Advisory Committee who funded the presentation of this paper at the regional Oklahoma-Arkansas Mathematical Association of America meeting.

References

Arkansas Vascular Flora Committee (AVFC). 2006. Checklist of the Vascular Plants of Arkansas. Arkansas Vascular Flora Committee, Fayetteville, Arkansas.

A. Cannon *et al., STAT2, Building Models for a World of Data,* Ed. (Freeman, New York, NY, 2013)

Biographical Sketches

Michael Lloyd graduated cum laude and in the honors program in Chemical Engineering with a B.S. in 1984. He accepted a position at Henderson State University in 1993 shortly after earning his Ph.D. in Mathematics (Probability Theory) from Kansas State University. He has presented papers at meetings of the Academy of Economics and Finance, the American Mathematical Society, the Arkansas Conference on Teaching, and the Southwest Arkansas Council of Teachers of Mathematics. He has been an active member of the Mathematical Association of America since 1993, earned 18 hours in computer science, and has been an Advanced Placement statistics consultant since 2002.

Jonathan Eagle received his B.S. in Biology, minoring in chemistry and statistics, in 2015 from Henderson State University. Graduating cum laude as member of Honors College and the McNair Scholar Program, he was recognized as the Outstanding Graduating Senior in the Biology Department. He plans to continue his education at the graduate level in the area of biomolecular sciences.

The Man-Forged Miscreants

Peter Wilson
Mentor: Peggy Dunn Bailey, Ph.D.

In this essay I deconstruct the facilitation we as people provide in the formulation of our most dangerous enemies. These miscreants are generally reflections of their creators, and often in literature they triumph over their creators in ironic or thought-provoking ways. To support this notion, I compare and contrast the antagonists from Mary Shelley's *Frankenstein* and Samuel Taylor Coleridge's *The Rime of the Ancient Mariner* using textual evidence and several critical responses.

Mary Shelley's *Frankenstein* utilizes a significant portion of its text deliberating what it means to be human. In the literal sense, Victor Frankenstein is the human and the creature a humanoid facsimile. Yet most readers identify Victor as the monster and his creation as a more emotionally human and relatable character. As the story progresses, it becomes clear to Victor what horror he has unleashed upon himself and his family. In denying the beast the fair treatment it craves, Victor creates his own arch nemesis.

In Samuel Taylor Coleridge's *The Rime of the Ancient Mariner*, this idea is similarly explored. The mariner's greatest obstacle throughout the text is divine retribution. His refusal