An eye-opener about ecology and mathematics

As much as I could during my undergraduate education, I avoided mathematics and statistics, and entered graduate school less well prepared to do modern research in ecology than many of the first year undergraduate students I would later teach. Ecology is a highly quantitative science, highly mathematical and statistical. This story will illustrate some of the consequences of that poor training, in hopes that others will not follow the shortsighted policy of avoiding subjects that are difficult for them.

The events I describe here occurred during the last year of my PhD studies, after I had left the field and was analyzing data on the determinants of territory size. It was shortly before my 32nd birthday. I had realized that hummingbird territory size was strongly dependent on flower density, but had not yet seen anything convincing in the data. The problem was that flower density, which was the independent variable in the relationship I needed to characterize, was not one variable but two - - four, actually, because territories contained two different species of flowers and they produced nectar at different rates. With a lot of help from my friends, I realized that I should combine all of the data about flowers into a single measure of the production of nectar energy by territories. I multiplied the number of flowers of each species in each territory by the rate at which that species produced nectar, then summed those values together into a synthetic variable and divided it by the area of the territory: nectar production density.

When I plotted 96 territory sizes against 96 nectar production densities, the points fell into a beautifully tight curve, swooping down from the large territories I had observed at low flower densities at the beginning and end of the season to the small ones in the best parts of meadows at the height of the season. I knew enough to realize that this graph was what I had been looking for, but not enough to characterize the curve statistically, which would be required for confirmation and for further analysis. I knew how to do a linear regression, but this relationship was not linear and I didn't have a clue what to do.

After fretting about this for a while, I took my graph to where Dan Udovic, a new ecology professor about my age, was teaching an undergraduate laboratory class. I caught his attention and motioned for him to come into the hallway. He took one look at the graph and said, "Do you realize what you have here?" I replied that I knew it was exciting and important, but didn't know what to do with it. Dan said he didn't have time to talk with me about it right then, but promised that if I performed a linear regression not on the territory sizes and nectar production densities directly, but on the logarithms of those variables, he would talk with me about it after his class was over.

In those days, all serious computer work used punched cards as input, and the cards were submitted at the computer centre in batches - - in this case, about 120 cards. I re-punched the card that specified the regression, submitted the job, and when Dan finished his class I was waiting for him. The statistical analysis was stunning in the clarity of the result that it provided. Whereas the plot of the raw data had been a swooping, concave-upward curve, the plot of the log-transformed data was a straight line with slope very near -1, indicating that all 96 territories were equivalent energetically. Every hummingbird in the system defended approximately the same amount of energy, regardless of the size of the territory, the density of flowers in the meadow, or anything else. Not only that, but the analysis showed that a large proportion of all the variation in territory size was explained by nectar production density alone. (Others later used more sensitive experimental approaches to determine that territory size is also affected by the number of competing hummingbirds; the more competition, the less nectar production territory holders can defend.)

What do I take from this story?

First, for Dan and my peers, who were unafraid of quantitative relationships and accustomed to thinking about them, it was an obvious thing to combine measures of flower density and nectar production rate to produce a synthetic variable. After all, the hypothesis I was testing was about partitioning of resources among competitors, and a moment's thought makes it obvious that the fundamental resource for hummingbirds is energy, not flowers or even nectar. I was blind to that without help, because my fear of statistics and mathematics clouded my vision. There was nothing at all complex about the operation, either; any 6th grader could have done it. All I had to do was multiply two numbers for each species of flowers, add those products together, and divide by the size of the territory. To perform the mathematical operations was not the issue, but to think creatively about the essence of a quantitative problem. I was woefully unprepared at the time to do that, even at that very basic level.

Second, it amazed me that in a couple of seconds, Dan Udovic completely understood the relationship my first graph depicted. Not only was he able to tell me how to perform the statistical analysis, but he already knew what the statistics would reveal and much of what I would be able to do with the relationship later. Much more profoundly in terms of my life, Dan understood in that brief moment that I would likely be able to use that one graph and what would follow from it to find a job in a research institution. But I was so afraid of mathematics and statistics that I didn't have more than the barest inkling of any of this.

What did Dan see in the graph?

Even before he had read the labels to see what variable I had plotted against what, he could tell that the plot was in the form of a hyperbolic function. Like the circle, the ellipse, the parabola, the straight line, and other basic functions, the hyperbola expresses graphically a unique kind of relationship between two variables, in this case territory size and nectar production density. If the relationship were indeed hyperbolic, its equation would reduce to energy units per territory, indicating the energetic equivalence of the territories.

These things were second nature to Dan, because he was accustomed to thinking about them. It was also obvious to him that if the plot of data on my graph were indeed a hyperbolic curve, plotting the logarithms of those data would transform the hyperbola into a straight line with slope of -1. And Dan knew that the statistics would evaluate how closely the relationship approximated a hyperbola and indicate what proportion of the variability in the data remained to be explained by other variables than those I had measured. It's all fairly basic stuff, but I was afraid of it.

I never lost my fear of mathematics. I did learn to stop allowing that fear to dictate my behaviour, however, and the revolution in computing that had not yet begun provided the means for me to rigorously explore quantitative aspects of the work without getting in too far over my head. And I learned to seek out experts to help me think about aspects of the work that were beyond me.

Dan Udovic is still a master at this kind of teaching. Thanks a lot, Dan!