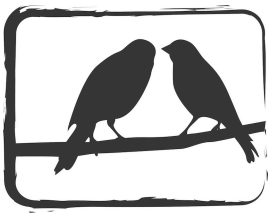
Environmental Humanities, vol. 7, 2015, pp. 59-88
www.environmentalhumanities.org
ISSN: 2201-1919



Mathematizing Nature's Messiness: Graphical Representations of Variation in Ecology, 1930-present

Laura J. Martin
Harvard University Center for the Environment, Harvard University, USA

ABSTRACT This article explores the rise of the line graph and an associated statistical method, linear regression, in ecology. At the turn of the 20th century, many ecologists studied variation in organismal traits, like height and weight, among populations of a species. The statistical practice of "polygons of variation" emerged out of such studies. But between 1930 and 1950, polygons of variation were gradually eclipsed by line graphs. Motivated by the recent and disastrous Dust Bowl, American ecologists began to place organismal variables and environmental variables on the same graph. They began to use linear regression, a then-obscure statistical method first developed to study whether parents passed their morphological traits onto offspring, to interpret these graphs. This use of linear regression marked an important shift in how ecologists interpreted biological variation. Variation—once ecologists' object of study—was now noise. Yet ecologists did not abandon their commitment to the idea that nature is complex, various, and interconnected. Rather, they came to read biologically meaningful patterns in seemingly "messy" graphs, using linear regression differently than other scientists, even those in closely allied disciplines. I situate this analysis in a broad STS literature on modeling that has tended to analyze the decision to model, or the choice of variables to include in a model, rather than how practitioners interpret the models they choose. I contend that not only has ecologists' use of linear regression shaped understandings of nature, but ecologists' understandings of nature have also shaped their use of linear regression.

Introduction

> It must be admitted that the ecologist is something of a chartered libertine; he roams at will over the preserves of the plant and animal biologists, the physiologist, the behaviorist, the meteorologist, the geologist, the physicist, the chemist, and even the sociologist. He poaches from all these and from other established and respected disciplines. It is indeed a major problem for the ecologist, in his own interest, to set bounds to his divagations.
> ~ Ayman Macfayden*, Animal Ecology: Aims and Methods*.[1]

> Ecology would be easy, were it not for all the ecosystems—vastly complex and variable as they are … Scientists like to impose structure and order on chaos, and ecologists are no different. Ecology has its grand theories, but they are riddled with conditional clauses,

[1] Ayman Macfayden, *Animal Ecology: Aims and Methods* (London: Pitman, 1963), xi.

> caveats and exceptions … It is doubtful that the generalities that underlie the complex patterns of nature will ever be phrased succinctly enough to fit on a T-shirt.
> ~ Editorial, *Nature*.[2]

From its inception at the turn of the 20th century, ecology has encompassed a wide range of practices. Today an ecologist might use radio tags to track the movement of Clark's Nutcrackers in the Rocky Mountains, or she might test the uptake of plant defensive chemicals by Monarch caterpillars in the laboratory. The discipline's intentional interdisciplinarity—its "poaching" from other disciplines, in the words of one classic textbook—has made it an attractive object of study for scholars interested in how scientific knowledge is produced.[3] Studies of ecology have led to wider insights about scientific negotiation, the workings of large formal knowledge networks, and how and when scientists involve themselves in environmental controversies.[4]

Most recently, scholars have turned to ecology to analyze how the places where scientists conduct experiments contribute to the credibility of scientific claims. In doing so, they have tended to characterize the laboratory and the field as foils. Laboratories—indoor spaces that can only be accessed by experts—appear to be sites of control, simplification, and mechanization. Field sites, meanwhile, appear to be sites of promiscuity, complexity, and imprecision, sites where nature has "free reign."[5] Robert Kohler, Sharon Kingsland, and other historians have argued that early ecologists sought to bring the prestige of laboratory science to ecology by incorporating laboratory instruments into their fieldwork.[6] In using thermometers, photometers, and other such instruments, ecologists addressed the pervasive concern that their discipline had become too easily accessible to amateur naturalists.

---

[2] Editorial, *Nature*, (13 March 2014): 139-140.

[3] Quote from Ayman Macfayden*, Animal Ecology: Aims and Methods* (London: Pittman, 1963), xi.

[4] See, for example, Susan Leigh Star and James R. Griesemer, "Institutional Ecology, 'Translations,' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39," *Social Studies of Science* 19 (1989): 387-420; Abby J. Kinchy and Daniel L. Kleinman, "Organizing Credibility: Discursive and Organizational Orthodoxy on the Borders of Ecology and Politics," *Social Studies of Science* 33 (2003): 869-896; Chunglin Kwa, "Local Ecologies and Global Science Discourses and Strategies of the International Geosphere-Biosphere Programme," *Social Studies of Science* 35 (2005): 923-950; Nicole L. Klenk, Gordon M. Hickey, and J. I. MacLellan, "Evaluating the Social Capital Accrued in Large Research Networks: The Case of the Sustainable Forest Management Network (1995-2000)," *Social Studies of Science* 40 (2010): 931-960; Myriah L. Cornwell and Lisa M. Campbell, "Co-producing Conservation and Knowledge: Citizen-based Sea Turtle Monitoring in North Carolina, USA," *Social Studies of Science* 42 (2012):101-120; F. Millerand, David Ribes, K. S. Baker, and G. C. Bowker, "Making an Issue out of a Standard: Storytelling Practices in a Scientific Community," *Science, Technology, & Human Values* 38 (2013): 7-43.

[5] D. W. Schneider, "Local Knowledge, Environmental Politics, and the Founding of Ecology in the United States: Stephen Forbes and 'The Lake as Microcosm' (1887)," *Isis* 91 (2000): 681-705; Jeremy Vetter, "Introduction," in *Knowing Global Environments: New Historical Perspectives on the Field Sciences* (New Brunswick: Rutgers University Press, 2011).

[6] Robert E. Kohler, *Landscapes and Labscapes: Exploring the Lab-Field Border in Biology* (Chicago: University of Chicago Press, 2002); Sharon Kingsland, *The Evolution of American Ecology, 1890-2000* (Baltimore: John Hopkins University Press, 2005).

Like Kohler's *Landscapes and Labscapes*, this article is concerned with how ecologists have negotiated the supposed tension between the "control" of the laboratory and the "unruliness" of the field. But rather than trace the use of instruments through time, it traces the use of a statistical practice: linear regression.[7]

Today linear regression is arguably the most widespread practice in ecology. However, ecologists took up linear regression recently relative to other practices like the use of quadrats. Indeed, in 1930 no ecologists used linear regression, even though economists, geographers, and other scientists had been using it since the mid-19th century. In the following section, I review the historical circumstances that led ecologists to adopt and adapt linear regression. I argue that the American Dust Bowl played a significant role in ecology's embrace of linear regression. Prior to the Dust Bowl, efforts to mathematize ecology were centered on studies of variation with single species. Into the 1930s, ecologists increasingly explored relationships between organismal variables (like plant productivity) and environmental variables (like rainfall). Linear regression was a means of considering variation in an organismal variable and an environmental variable simultaneously.

But as ecologists incorporated linear regression into their practice, they came to use it differently than scientists of other disciplines, as the second section of the article explores. Ecology's "messy graphs" say more about the history of ecology than they do about the workings of the natural world. Not only has linear regression shaped ecological understandings of nature, but ecological understandings of nature have also shaped linear regression. I situate this analysis in a broad STS literature on modeling that has tended to analyze the decision to model, or the choice of variables to include in a model, rather than how practitioners interpret the models they choose.

Robert McArthur began his textbook *Geographical Ecology* with a stern reprimand: "Not all naturalists want to do science," he wrote, "many take refuge in nature's complexity as a justification to oppose any search for patterns."[8] Here I contend that ecology's commitment to biological complexity has not hindered its search for pattern. Rather, through linear regression, ecologists have successfully established the idea that nature is various, messy, and complex, *while simultaneously* establishing the idea that ecologists are the experts on how to see pattern in this complexity.

From Polygons of Variation to Linear Regression

In the 1930s the American Great Plains suffered some of the worst droughts in its history, displacing settlers across the region.[9] When this "Dust Bowl" began, ecology was not the well-

---

[7] Since Shapin and Schaffer (1985), STS scholars have devoted a marked attention to objects, using them to explain broader ideas and processes, as in the case of St. Brieuc scallops (Callon, 1986), a hotel key (Latour, 1991), the cloud chamber (Galison, 1997), and specimens at the Berkeley Museum of Vertebrate Zoology (Star and Griesemer, 1989). Here I draw attention to the assumptions that shape statistical practices, which, in turn, shape understandings of material nature.

[8] Robert H. MacArthur, *Geographical Ecology: Patterns in the Distribution of Species* (New York: Harper & Row, 1972).

[9] For cultural histories of the American Dust Bowl, see Donald Worster, *Dust Bowl: The Southern Plains in the 1930s* (New York: Oxford University Press, 1979); The collection of essays in *Great Plains Quarterly* 6 (1986); James Gregory, *American Exodus: The Dust Bowl Migration and Okie Culture in*

organized, influential scientific discipline it is today.[10] The Ecological Society of America (ESA) had been founded in 1915 by a group of botanists and zoologists interested in questions at the intersection of their disciplines. The British Ecological Society had been founded two years prior. By 1930, the ESA—today an organization of more than 10,000 members—had about 500 members, and very few of these members referred to themselves as "ecologists," identifying instead as botanists, zoologists, or entomologists.[11]

Biometry, however, was a thriving subfield. Darwin's theory of natural selection had spurred a lasting interest in organismal variation. At the turn of the century, zoologists began applying "mathematical statistics"—recently developed in Europe as municipalities systematized the collection of demographic and economic data to study characteristics of human populations such as the average number of crimes per year—to species other than humans.[12] The first of these studies was in published in 1889 by Raphael Weldon, a British zoologist. In "The Variations Occurring in Certain Decapod Crustacea," Weldon graphed the distribution of four morphological traits in a shrimp population. By comparing the distribution of morphological variation among populations, Weldon then attempted to test popular biogeographical theories like the theory that animals in colder regions had evolved to be larger than those in warmer regions.[13]

California (New York: Oxford University Press, 1989); Jani Scandura, *Down in the Dumps: Place, Modernity, American Depression* (Duke University Press, 2008).

[10] On early American ecology and connections to European biology and geography, see Donald Worster, *Nature's Economy: A History of Ecological Ideas* (Cambridge: Cambridge University Press, 1977), Chapters 1-9; Janet Browne, *The Secular Ark: Studies in the History of Biogeography* (New Haven: Yale University Press, 1983); Eugene Cittadino, *Nature as the Laboratory: Darwinian Plant Ecology in the German Empire, 1880-1900* (Cambridge: Cambridge University Press, 1990); Lynn K. Nyhart, *Biology Takes Form: Animal Morphology and the German Universities, 1800-1900* (Chicago: University of Chicago Press, 1995); Peder Anker, *Imperial Ecology: Environmental Order in the British Empire, 1895-1945* (Cambridge: Harvard University Press, 2002); Aaron Sachs, *The Humboldt Current: Nineteenth Century Exploration and the Roots of American Environmentalism* (New York: Viking, 2006). For ecologists' accounts of their early disciplinary history, see: Victor E. Shelford, "The Organization of the Ecological Society of America 1914-19," *Ecology* 19 (1938): 164-166; Arthur G. Tansley, "The Early History of Modern Plant Ecology in Britain," *Journal of Ecology* 35 (1947): 130-137; Norman Taylor, "The Beginnings of Ecology," *Ecology* 19 (1938): 352.

[11] Robert Burgess, *Historical Data and Some Preliminary Analyses* (Washington D.C.: Ecological Society of America, c. 1976).

[12] Theodore M. Porter, *The Rise of Statistical Thinking, 1820-1900* (Princeton: Princeton University Press, 1986); Ian Hacking, *The Taming of Chance* (Cambridge: Cambridge University Press, 1990); Theodore M. Porter, *Trust In Numbers: The Pursuit of Objectivity in Science and Public Life* (Princeton: Princeton University Press, 1995); Anders Hald, *A History of Parametric Statistical Inference from Bernoulli to Fischer, 1713-1935* (New York: Springer Science, 2007).

[13] R. Weldon, "The Variations Occurring in Certain Decapod Crustacea I. *Cragon vulgaris*," *Proceedings of the Royal Society* 47 (1889): 286-291.

Debates over heredity and genetics further heightened interest in studies of organismal variation.[14] In 1901, Weldon and Karl Pearson, a professor of applied mathematics at the University College London, founded the journal *Biometrika*.[15] In the first issue, Weldon and Pearson explained that while the starting point of Darwin's theories was the existence of variation among individuals of a species, variation could not be "an effective factor in evolution" unless it appeared across many individuals. The study of evolution, they continued, demanded statistical analysis:

> It is not a mere formal clothing of biological conceptions with mathematical symbols that is here indicated, or that we are considering, when we say that all Darwin's ideas fit themselves to algebraic definition. On the contrary—exactly as in the like case of the mathematical treatment of Faraday's conceptions of electromagnetism—the symbolic analysis widens our notions, it leads us at once to new points of view and it directly suggests—perhaps this is its most important advantage—fresh points for observation and novel directions for experimental research.[16]

Statistical analysis, Weldon and Pearson concluded, would provide biologists with a new tool for exploring the theories of 19th century naturalists.[17]

In the first years of *Biometrika*'s publication, the most popular statistical method among its authors was the construction of "polygons of variation." A polygon of variation displayed a morphological trait, like height or weight, against its frequency—similar to what we refer to as a histogram today. By comparing the polygons of variation of populations from different environments, biometricians hoped to reveal patterns of evolutionary change. Francis Galton extolled biometry's promise of revealing pattern in biological data, of "converting a mob into an orderly array."[18]

---

[14] e.g., Francis Galton, "Co-relations and their Measurement, Chiefly from Anthropometric Data," *Proceedings of the Royal Society of London* 45 (1888): 135-145; Francis Galton, *Natural Inheritance* (New York: Macmillan and Company, 1894).

[15] D. R. Cox, "Biometrika: The First 100 Years," *Biometrika* 88 (2011): 3-11.

[16] See *Biometrika* 1 (1901): 1-6.

[17] Camic and Xie (1994) argue that turn-of-the-century anthropologists, sociologists, psychologists, and economists were doing boundary work to legitimize their disciplines in "a competitive interdisciplinary field" when they turned to statistical methods to "demonstrate compliance with acceptable scientific models and at the same time carve out a distinctive mode of statistical analysis to differentiate their own discipline from the others." This article explores the emergence of ecology's distinctive mode of linear regression. Charles C. Camic and Yu Xie, "The Statistical Turn in American Social Science: Columbia University, 1890 to 1915," *American Sociological Review* 59 (1994): 773-805.

[18] Francis Galton, "Biometry," *Biometrika* 1 (1901): 7-10, 7.

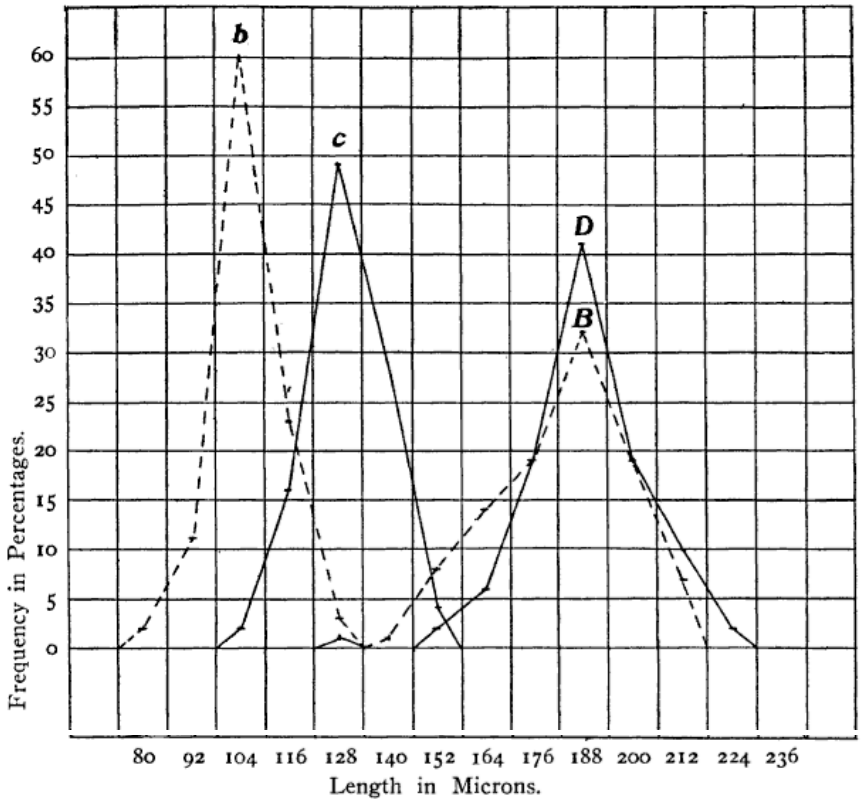


Figure 1. *A "polygon of variation," or histogram, from Jennings (1908). The graph displays the frequency (y-axis) of length (x-axis) in four descendent populations of paramecia. Curve b, for example, conveys that 60% of individuals from population b had a length of 110 microns. B is the polygon for 100 descendants of ten larger individuals. D is the polygon for 100 descendants of a single large individual. b is the polygon for 100 descendants of 50 smaller individuals. c is the polygon for 100 descendants of a single small individual.*

At the time that mathematicians and zoologists were organizing biometry into a discipline, a handful of European and American biologists had begun to identify as ecologists. Many early members of the British Ecological Society and the Ecological Society of America were zoologists trained in biometry. Among these, Moore, Pearson, and others worked to introduce biometric methods to other ecologists. For instance, in 1920 zoologist Ellis L. Michael published "A Plea in behalf of Quantitative Biology" in the BES's newly established *Journal of Ecology*, arguing that soon statistical fluency would be as important to biologists as language fluency. But such positions met some resistance. The editors of *Journal of Ecology* published Michael's piece, but with a footnoted caveat: "We are glad to publish Mr. Michael's plea. But we may question if it is practicable to insist on 'proficiency in mathematics' for *all* biologists at this stage of the development of science."[19]

Such exchanges were frequent in ecology journals through the 1920s, as scientists from wildly different backgrounds debated the appropriate balance between descriptive and mathematical work. Robert Kohler has contended that early ecologists sought to mathematize

[19] E. L. Michael, "Marine Ecology and the Coefficient of Association: A Plea in Behalf of Quantitative Biology," *Journal of Ecology* 8 (1920): 54-59, 59.

ecology and "become physiologists of the field" in order to make the discipline inaccessible to naturalist hobbyists and amateur botanists.[20] Henry Cowles, for example, a prominent botanist at the University of Chicago, lamented the "many 'contributions' to ecology which consist of a hasty gathering together of notes made in leisure moments during summer holidays."[21]

Biometric statistics were one means of making ecological studies more specialized. Through the 1920s some ecologists adopted biometric methods. But ecologists were specifically interested in the interaction between organisms and their environments. Polygons of variation captured only variation within populations of organisms—variables like length, height, or weight. In 1930, Russian zoologist G. F. Gause observed:

> As is well known, these last years have been marked by great progress in descriptive ecology, considered in the broadest sense of the term. Nevertheless a whole range of questions in this field have so far been but little investigated. We refer to problems bearing upon the exact study of the distribution of organisms in its relation to the factors of environment.[22]

It should be possible, Gause continued, to place environmental factors like temperature or relative humidity on one axis of a graph and the abundance of a species on another. The resultant graphs would "establish the average ecological conditions for this or that species."[23] They would be polygons of variation, but with environmental variables rather than organismal variables on the x-axis. In this way, one could display how variation in an environmental factor like temperature related to the distribution of, say, different populations of Paramecium, or, as Gause would become famous for with his "competitive exclusion principle," different species entirely.

Thus polygons of variation provided a means of displaying organismal variation *or* environmental variation. But not both together. Graphical representations of organismal variation *and* environmental variation came later, with line graphs. Beginning in the late 1700s, geographers and economists began constructing line graphs to study the relationship between human population and economic variables.[24] The first person to use line graphs to explore a biological question was Francis Galton, English polymath and cousin of Darwin. In a study comparing the heights of parent sweet peas with those of their offspring, Galton articulated the idea of "co-relation" (correlation). He wrote: "two variable organs are said to be correlated when the variation of the one is accompanied on the average by more or less variation of the

---

[20] Kohler, *Landscapes and Labscapes*, 96.

[21] Henry Cowles, "Review of Research Methods in Ecology by F. E. Clements," *Botanical Gazette*, 40 (1905): 381-382.

[22] G. F. Gause, "Studies on the Ecology of the Orthoptera," *Ecology* 11 (1930): 307-325, 307. Gause and other ecologists were likely exposed to Pearson's work on correlation through George Udny Yule's 1927 edition of *An Introduction to the Theory of Statistics* (London: C. Griffin & Co., 1927).

[23] Gause, "Studies on the Ecology of the Orthoptera," 308.

[24] – Scottish engineer William Playfair first developed the line graph (along with the pie chart and the bar chart) in his 1786 *Commercial and Political Atlas*. William Playfair, *The Commercial and Political Atlas* (London, 1786).

other, and in the same direction."[25] In his work on heredity, Galton developed mathematical methods for summarizing the correlation between two variables.

Galton's work would remain obscure until Pearson reviewed it in a 1920 *Biometrika* article.[26] Through "Notes on the History of Correlation" and subsequent articles, Pearson introduced ecologists to a century of work on line graphs and the visualization of correlation. He also expanded upon Galton's mathematical treatment of correlation and "linear regression"—fitting a straight line through a set of points of a data set in such a way as to make vertical distances between the points and the fitted line as small as possible (Figure 2). Soon, a few biometrists and ecologists like G. F. Gause began promoting the use of linear regression.[27]

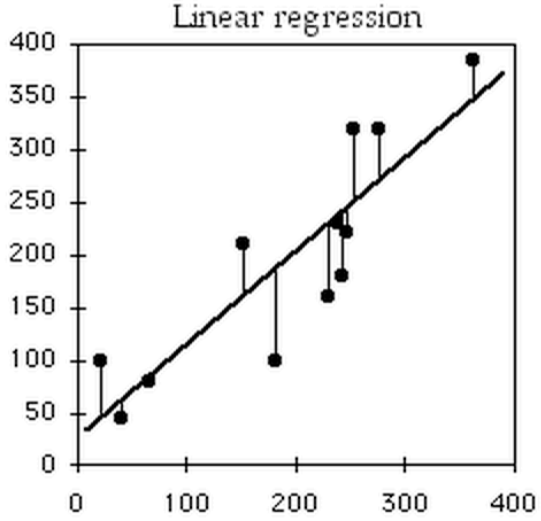


Figure 2. *Linear regression finds the line that "best fits" all of the data points. This graph shows the data points (dots), linear regression line (thick line), and the distances between the data points and the regression line (thin lines). The regression line is the line that minimizes the sum of the squared vertical distances between the points and the line. From McDonald (2014). Reproduced with permission.*

But it was the American Dust Bowl that solidified linear regression's popularity. From 1930 to 1934, severe droughts led to dust storms that forced tens of thousands of families to abandon their farms on the U.S. and Canadian prairies. Plowed soil turned to dust, and winds blew that dust into devastating clouds, "black blizzards" and "black rollers" that reached as far east as

[25] Galton, "Co-relations and their Measurement," 135.

[26] Karl Pearson, "Notes on the History of Correlation," *Biometrika* 13 (1920): 25-45; Karl Pearson, *Francis Galton: A Centenary Appreciation* (Cambridge: Cambridge University Press, 1922); S. M. Stigler, *The History of Statistics: The Measurement of Uncertainty Before 1900* (Cambridge: Harvard University Press, 1986); M. D. Friendly, "The Early Origins and Development of the Scatterplot," *Journal of the History of the Behavioral Sciences* 41 (2005):103-130.

[27] Porter, *The Rise of Statistical Thinking*; S. M. Stigler, "Francis Galton's Account of the Invention of Correlation," *Statistical Science* 4 (1989): 73-79; Stephen Blyth, "Karl Pearson and the Correlation Curve," *International Statistical Review* 62 (1994): 393-403.

New York City and Washington, D.C.[28] In his first days in office, President Franklin Delano Roosevelt instituted the New Deal to address the combined natural and national disaster. New Deal programs, especially the Civilian Conservation Corps, drastically reorganized federal land management, and beginning in 1934 the federal government purchased 11 million acres of "exhausted" land from private owners. As a result of these reforms, the federal government acquired ownership or control of a large amount of new, marginal land.[29]

Ecologists soon recognized the professional opportunities the New Deal afforded, and they worked to fashion themselves as experts on land management. In "Experimental Ecology in the Public Service," for example, botanist Frederic Clements explained that the "tragic process" that had led to the Dust Bowl would continue until citizens recognized the importance of ecological research.[30] Addressing the Ecological Society of America in St. Louis in 1935, ESA president Walter P. Taylor asserted that the solution to wasteful cropping practices, overgrazing, and marshland draining lay in ecology: "Who but a geo-bio-ecologist, one who knows something of interrelationships and of plant and animal indicators and soils, is qualified for the important tasks of land classification?"[31]

The Dust Bowl's particular crisis spurred an interest in the factors controlling plant growth and soil erosion and led to exchange between ecology and agronomy. While some agronomists argued that crop yields could be improved by increased irrigation, others believed temperature was the most important factor controlling plant growth.[32] Such debates invited studies expressing the relationship between two or more varying parameters. Among American agronomists, linear regression gained traction after the 1934 and 1936 U.S. Department of Agriculture sponsored statistical conferences at Iowa State University. And into the 1930s, an increasing number of ecologists began to employ linear regression to study the relationship between environmental variables (like rainfall) and biological variables (like plant productivity). In *Biometrika*, *Biometrics*, *Journal of the American Statistical Association*, and *Journal of the Royal Statistical Society*, statisticians and ecologists began to expand upon linear regression methods.

Unsurprisingly, precipitation and temperature were among the first environmental variables that ecologists attempted to graph in relation to plant growth. Court (1930) investigated the relationship between plant growth, temperature, and soil moisture. "High temperatures that would stimulate in the presence of adequate water might mean death in a drought," he wrote.[33] Hawley (1937) asked whether annual growth ring width in red cedar was correlated with precipitation, using the Pearson correlation coefficient to relate the mean ring-

---

[28] Donald Worster, *Dust Bowl: The Southern Plains in the 1930s* (New York: Oxford: Oxford University Press, 1979); Alexander J. Field, *A Great Leap Forward: 1930s Depression and U.S. Economic Growth* (New Haven: Yale University Press, 2011).

[29] Neil Maher, *Nature's New Deal: The Civilian Conservation Corps and the Roots of the American Environmental Movement* (New York: Oxford University Press, 2008).

[30] Frederic E. Clements, "Experimental Ecology in the Public Service," *Ecology* 16 (1935): 342-363.

[31] Walter P. Taylor, "What is Ecology and What Good is it?" *Ecology* 17 (1936): 333-346, 340.

[32] Thank you to David Alan Grier for this information. See also George Snedecor's 1938 Statistical Methods book.

[33] Andrew T. Court, "Measuring Joint Causation," *Journal of the American Statistical Association* 25 (1930): 245-254.

size for each year to the hydrologic data.[34] Diller (1935) observed that in certain regions, temperature seemed to have the greatest effect on the distribution and growth of forest trees, where elsewhere precipitation seemed to be the most important variable (Figure 3).[35] Recent drought emphasized "the need of more exact knowledge of the factors influencing survival of trees," wrote Shirley and Meuli (1939) in an article presenting the results of linear regression between soil nutrients, soil moisture, and Red Pine growth.[36]

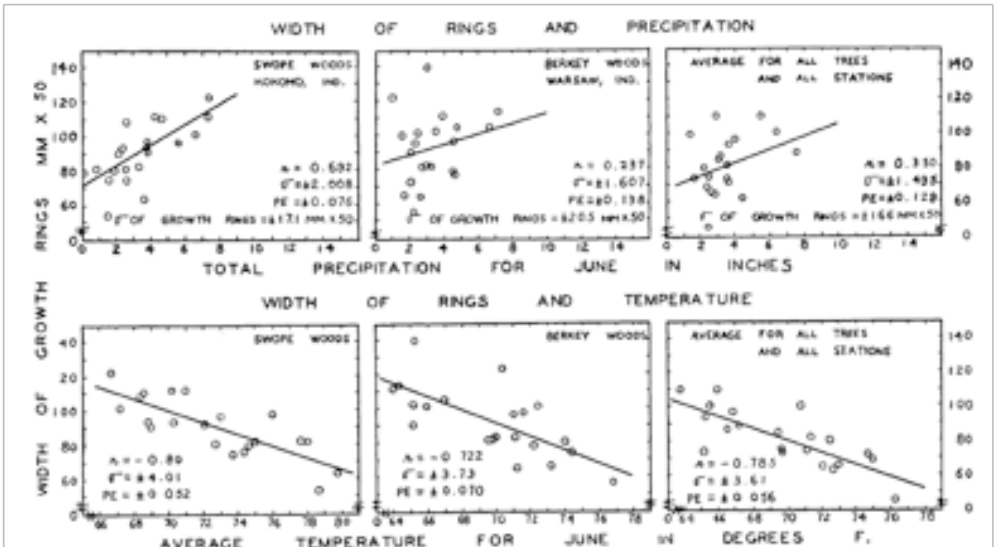


Figure 3. *Line graphs from Diller (1935) depicting the correlation of width of beech growth rings (y-axis) with precipitation (x-axis, top) and temperature (x-axis, bottom) for samples from Swope woods (left panels), Berkey woods (middle panels), and their averages (right panels).*

Through such studies regression slowly gained traction in ecology, so that by the 1950s many ecologists were familiar with the statistical method. And ecologists' uses of linear regression were importantly different than those of other scientists. Economists, chemists, and others used linear regression to derive estimates or to extend trends into the future. For example, Williams

[34] F. M. Hawley, "Relationship of Southern Cedar Growth to Precipitation and Run Off," *Ecology* 18 (1937): 398-405.

[35] Oliver D. Diller, "The Relation of Temperature and Precipitation to the Growth of Beech in Northern Indiana," *Ecology* 16 (1935): 72-81.

[36] Hardy L. Shirley and Lloyd J. Meuli, "The Influence of Soil Nutrients on Drought Resistance of Two-year-old Red Pine," *American Journal of Botany* 26 (1939): 355-360. Agronomists also used linear regression to study the relationship between environmental factors and plant growth. For example, the "Crop-Weather-Yield Project" was set up by the Agricultural Marketing Service, cooperating with various State Agricultural Experiment Stations, to determine the physiological effects of climate on plant development. This work was motivated by a desire to predict future corn crops based on weather variables. The resources at stake were substantial: dry weather reduced corn yields by more than 50%, Shaw and Loomis (1950) reported. R. H. Shaw and W. E. Loomis, "Bases for the Prediction of Corn Yields," *Plant Physiology* 25 (1950): 225-244. See also F. E. Davis and G. D. Harrell, *Relation of the Weather and its Distribution to Corn Yields*. USDA Technical Bulletin 806 (Washington DC: Government Printing Office, 1942).

(1959) presented linear regression as an accounting shortcut for the timber industry. Rather than take data on the cost of individual logs, a company could graph the total daily cost against the total number of logs in each size class to estimate the cost of individual logs.[37]

Ecologists, meanwhile, used linear regression towards another end: to determine *whether* there was a relationship among two or more variables. Schultz (1956) explained that there were "two kinds of tools" that ecologists could "carry with them" to the field: mechanical tools and statistical tools. Mechanical tools included meter sticks, nets, and axes. Statistical tools included linear regression, a "powerful tool for the analysis of data after the measurements have been taken." Unlike laboratory scientists, Schultz argued, ecologists had to contend with complex natural environments:

> Plant physiologists who are bequeathed with unlimited funds have elaborate laboratories and greenhouses where nearly every essential feature of the environment can be controlled. Thus, an experiment can be reduced to only one variable such as growth. With complete control over all factors, there should, theoretically, be no unexplained error encountered in the experimentation. […] Ecologists have two strikes against them—they never are bequeathed with unlimited funds and if they were, they would fall short in controlling most factors of the outdoor environment, as the rainmakers can attest. So their research is redolent with what is called experimental error.[38]

But, Schultz continued, linear regression provided ecologists with a tool to account for environmental variation. While ecologists could not control environmental variables, they could measure them:

> In fact, the ecologist may be better off than the physiologist because in many cases statistical control is more desirable than experimental control. First, the actual situation is studied, not one produced artificially; second, a far greater range of observation can be made which broadens the foundation for inference; and finally, one learns how two quantities instead of one vary, singly and together.[39]

Linear regression provided an objective method of analyzing the contribution of multiple environmental factors to organismal variation, Schultz concluded. "If these factors are real, they are measurable; and, if measurable, they are interpretable."[40]

In 1959, Australian statistician E.J. Williams published *Regression Analysis*, a textbook that enjoyed wide popularity among ecologists. In the preface of *Regression Analysis*, Williams argued that a treatise dealing with the relations among two or more variables was long overdue. Regression analysis had proven to be most useful in the biological sciences, he continued, wherein "the idea of a relationship among errorless quantities turns out to be otiose."[41] The following year, botanist C. Wayne Cook published a review of previous uses of linear

---

[37] E. J. Williams, *Regression Analysis* (New York: John Wiley and Sons Inc, 1959).

[38] Arnold Schultz, "The Use of Regression in Range Research," *Journal of Range Management* 9 (1956): 41-46.

[39] Arnold Schultz, "The Use of Regression in Range Research," *Journal of Range Management* 9 (1956): 41-46.

[40] Ibid.

[41] Williams, *Regression Analysis*.

regression in ecology. Most of these articles explored the relation between soil moisture and plant growth.[42] Wayne's own work explored the relationship between soil moisture and the sagebrush abundance in the Great Basin. Ecologists had applied linear regression, a technique that biometricians had developed to study heredity, to an entirely new type of inquiry. Biometricians had used linear regression to compare parents to their offspring. Ecologists were using linear regression to compare organisms to their environments.

From the 1930s to around 1970, it was mostly plant ecologists using linear regression, a method first introduced to ecology by zoologists studying heredity. In the 1970s linear regression was taken up again by animal ecologists, though this time to study the relationship among organismal and environmental variables. James (1970) used linear regression to explore the relationship between Downy Woodpecker wing length and air temperature (Figure 4).[43] With linear regression, Grant (1971) analyzed the relationship between deer mouse population size and the number of grassland plant species present.[44] The increasing availability of computers allowed ecologists to include an ever greater number of variables in regression analyses. Pugesek and Diem (1983), for example, tested for relationships among seagull offspring mortality and parental age, nest location, habitat, and clutch size, using the Bowling Green State University IBM computer.[45]

[42] C. Wayne Cook, "The Use of Multiple Regression and Correlation in Biological Investigations," *Ecology* 41 (1960): 556-560. See also W. T. Edmondson, "Reproductive Rate of Planktonic Rotifers as Related to Food and Temperature in Nature," *Ecological Monographs* 35 (1965): 61-111; Norman R. Draper and H. Smith, *Applied Regression Analysis* (New York: John Wiley and Sons Inc., 1966); G. A. Yarranton, "Plant Ecology: A Unifying Model," *Journal of Ecology* 57 (1969): 245-50; R. Mead, "A Note on the Use and Misuse of Regression Models in Ecology," *Journal of Ecology* 59 (1971): 215-219; A. O. Nicholls and D. M. Calder, "Comments on the Use of Regression Analysis for the Study of Plant Growth," *New Phytologist* 72 (1973): 571-581.

[43] Frances C. James, "Geographic Size in Birds and its Relationship to Climate," *Ecology* 51 (1970): 365-390.

[44] P. R. Grant, "Experimental Studies of Competitive Interaction in a Two-Species System," *Journal of Animal Ecology* 40 (1971): 323-350.

[45] Bruce H. Pugesek and Kenneth L. Diem, "A Multivariate Study of the Relationship of Parental Age to Reproductive Success in California Gulls," *Ecology* 64 (1983): 829-839.

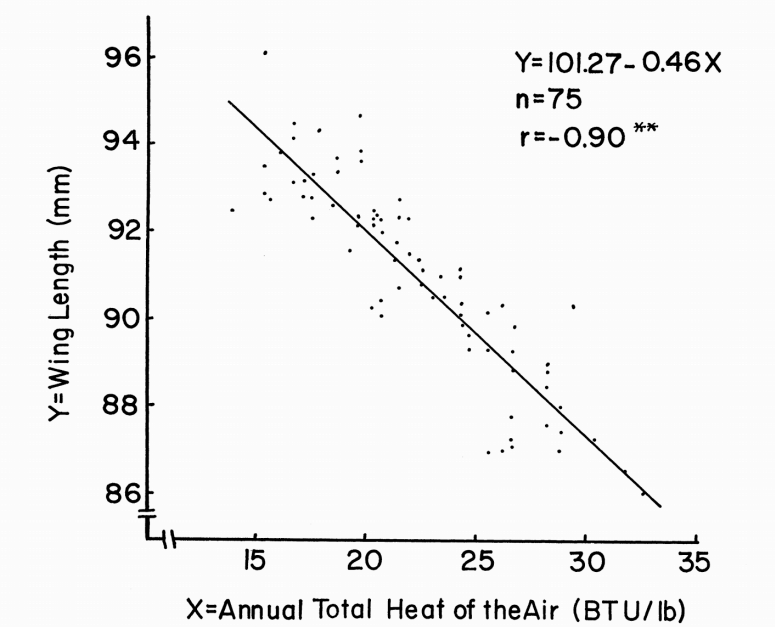


Figure 4. *A line graph from James (1970) depicting a linear regression of Downy Woodpecker wing lengths (y-axis) and air temperature (x-axis).*

In the inaugural issue of *Biometrika*, Weldon and Pearson (1901) circumspectly wrote that in adopting statistical practices, biologists faced the "danger" that "mathematics may tend to diverge too widely from Nature." Mathematics was abstract and orderly, they contended, whereas nature was concrete and messy. Little did they know that, a century later, it would be impossible to publish an article in an ecology journal that did not include statistical analysis. While in 1930 virtually no articles published in *Ecology* or *Journal of Ecology* reported statistics, by approximately 1975, half of them did, and by 1980 more than 75% did.[46]

But by no means did ecologists surrender their belief in nature's complexity. The following section explores in more detail how ecologists came to employ linear regression differently than scientists of other disciplines, seeing relationships and even causality in what we might call "messy" graphs. Not only did ecologists' use of linear regression shape their understandings of nature, but their understandings of nature also shaped their use of linear regression. Ecologists expected relationships among organismal variables and environmental variables to be difficult to perceive. Their use of linear regression both depended on and reproduced the idea that nature is complex, unruly, and entangled, yet knowable.

Ecology and "Messy" Graphs

It is rare that the data in ecological graphs fall perfectly along a line. Consider Figures 3 and 4. The scatter of data points around the regression lines are substantial. Indeed, today ecology is known for its complex or "messy" data and is often critiqued as a "soft science" on this basis. The view is summarized by the title of a recent article in *Frontiers in Ecology and the Environment*: "Rising Complexity and Falling Explanatory Power in Ecology."[47]

---

[46] Low-Decarie, Etienne, Corey Chivers, and Monica Granados, "Rising Complexity and Falling Explanatory Power in Ecology," *Frontiers in Ecology and the Environment* 12 (2014): 412-418.

[47] Low-Decarie *et al.,* (2014)

Just as the incorporation of linear regression into ecological methods was historically contingent, so, too, was the way in which ecologists *employed* linear regression. For linear regression was not and is not a stable or one-dimensional thing (excuse the pun). Where some scientists see pattern, others see none.

The "messiness" of a graph can be summarized by a single number, the "R^2 value" (also known as the coefficient of determination). The R^2 value is the proportion of variation in the dependent variable (aka "Y," "response variable," or "regressand") that can be attributed to variation in the independent variable(s) (aka "X," "explanatory variable," or "regressor"). In other words, R^2 is the proportion of variability in a data set that the statistical model accounts for. It is a summary of the spread of points around a regression line—one number that conveys the "messiness" of a graph.[48] An R^2 value of 1 indicates that the data points fall perfectly along the regression line. An R^2 value of 0 indicates that there is no relationship between the two (or more) variables (Figure 5). Thus "messy" graphs have low R^2 values.

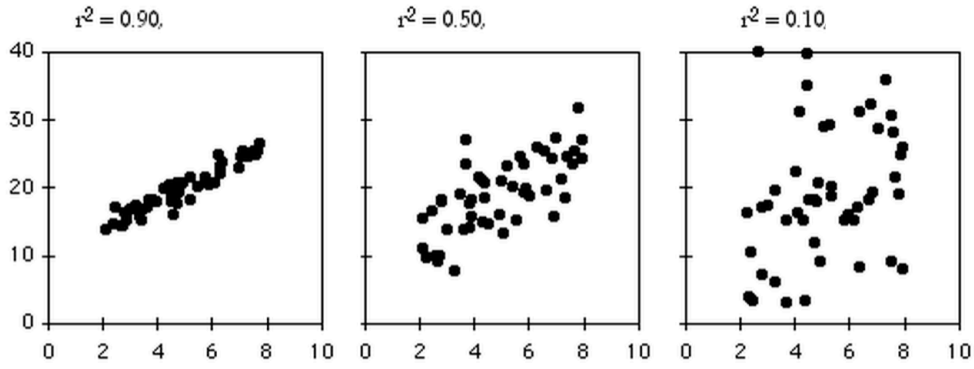


Figure 5. *Three relationships with the same slope, same intercept, and different amounts of scatter around the regression line, hence different R values from high—left, to low—right. Figure adapted from McDonald (2014). Reproduced with permission.*

What can R^2 values reveal about the history and philosophy of ecology? Studies of another statistical value, the p-value (aka "significance value"), have importantly revealed a bias in scientific literature towards "positive results." The p-value is used in null hypothesis testing to quantify the idea of statistical significance of evidence. If the p-value is less than a pre-set threshold value (historically 5% or 1%), the researcher rejects the null hypothesis and accepts the alternative hypothesis. Science studies scholars have demonstrated that, in a world where scientists are increasingly evaluated on the number of citations they receive, scientists are less likely to submit studies for publication that do not reject the null hypothesis.[49] Moreover,

[48] In recent years linear regression has been extended to generalized linear regression, which can accommodate many different types of response variables, and to the general linear model, which allows for multivariate response data that unifies commonly used methods including t-test, ANOVA, ANCOVA, and redundancy analysis.

[49] R. Rosenthal, "The File Drawer Problem and Tolerance for Null Results," *Psychological Bulletin* 86 (1979): 638-641.

journals are less likely to publish studies that do not reject the null hypothesis.[50] These two forces have contributed to a positive-outcome bias across many scientific literatures. In ecology, for example, approximately 90% of contemporary studies report statistically significant results.[51]

Like p-values, R^2 values can illuminate how scientists negotiate disciplinary norms and evaluate evidence. In ecology, economics, and sociology, the R^2 value is often interpreted as a measure of the influence of an independent variable on the dependent variable.[52] But, unlike the case for p-values, there is no R^2 value at which ecologists explicitly deem a model invalid or even implausible. Ecologists will see a relationship between two variables in a graph even when there is a lot of scatter, or a very low R^2 value.

In an automated analysis of 18,076 articles published from 1930 to 2010 in three ecology journals, Low-Decarie *et al.* found an average R^2 value of 0.55.[53] In other words, in a typical article, approximately 55% of variation in the dependent variable was "explained by" variation in the independent variable(s). An R^2 value of 0.55 is low, considering that many physical sciences and some medical subfields routinely report R^2 values of 0.99. Even closely allied fields, like biochemistry and climate science, report higher R^2 values. In the interdisciplinary journal *Science* in the year 2000, the mean R^2 value for ecology articles was 0.51, while the mean R^2 value for climate science articles was 0.77 (Figure 6).[54]

---

[50] P. J. Easterbrook, R. Goplan, J. A. Berlin, and D. R. Matthews, "Publication Bias in Clinical Research," *The Lancet* 337 (1991): 867-872; T. D. Sterling, W. L. Rosenbaum, and J. J. Weinkam, "Publication Decisions Revisited—The Effect of the Outcome of Statistical Tests on the Decision to Publish and Vice Versa," *American Statistician* 49 (1995): 108-112.

[51] R. D. Csada, P. C. James, and R. Espie. "The "File Drawer Problem" of Non-significant Results: Does it Apply to Biological Research?" *Oikos* 76 (1996): 591–593; Daniele Fanelli, "Negative Results are Disappearing from Most Disciplines and Countries," *Scientometrics* 90 (2012): 891-904. This phenomenon has led some ecologists and statisticians to critique the "overemphasis" of significance testing, and specifically, that the value of 0.05 has become the "absolute limit between two worlds" N. G. Yoccoz, "Use, Overuse, and Misuse of Significance Tests in Evolutionary Biology and Ecology," *Bulletin of the Ecological Society of America* 72 (1991): 106-111.

[52] F. Filho, J. A. Silva, and E. Rocha, "What is R^2 All About?" *Leviathan* 3 (2011): 60-68.

[53] Low-Decarie, Etienne, Corey Chivers, and Monica Granados, "Rising Complexity and Falling Explanatory Power in Ecology," *Frontiers in Ecology and the Environment* 12 (2014): 412-418.

[54] To analyze this I sampled 20 articles randomly from the year 2010 in *Ecology* and *Science*. This, of course, is a back-of-the-envelope calculation, and a larger sample size would be needed to make fine-scale comparisons between disciplines. I chose to consider climate science because, like ecology, climate science strives to interpret environmental variables. My impression from reviewing biochemical and medical journals is that R^2 values are also higher in those disciplines than in ecology. I speculate that R^2 values in sociology and economics are comparable to those in ecology. Møller and Jennions (2002), who also attempt to quantify mean R^2 values in ecology articles, report an average of 0.025—one twentieth of the value I found in this limited review, and those reported by Low-Decarie *et al*. A. P. Møller and M. D. Jennions, "How Much Variance can be Explained by Ecologists and Evolutionary Biologists?" *Oecologia* 132 (2002): 492-500.

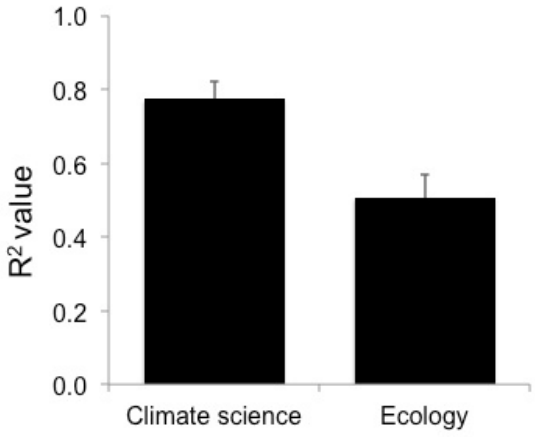


Figure 6. *The average R² values in articles reporting positive results in the journal Science in 2000.*

As outlined in the previous section, ecologists' interest in linear regression first stemmed from climatological interests. Why then do the R^2 values of contemporary ecological models and climate models differ? And why has the average R^2 value reported in ecology articles hovered around 0.55 since widespread use of linear regression began, even though many of the methods, tools, and emphases of ecology have changed?

Many argue that biological variables are inherently "messier" than physical variables. Indeed, many ecologists have lamented their inability to produce "grand laws."[55] For instance, Lindenmayer and Hunter have suggested that while physics, chemistry, and mathematics have "laws that form the backbone of those disciplines," the search for generalities in ecology is "thwarted by contingency and ecological complexity that limit the development of predictive rules."[56] I would contend, however, that these low R^2 values reveal more about how ecologists see the natural world than about the structure of the natural world itself.

Indeed, the idea that organismal and environmental variables are inherently various cannot be found in the methods sections of academic articles that employ linear regression. Ecologists have not felt compelled to justify their acceptance of messy graphs by appeals to the nature of their object of study. Rather, the idea has been articulated elsewhere, in review articles, textbooks, and in the precedents set by previous ecologists. For instance, in their review article, "How Much Variance can be Explained by Ecologists and Evolutionary Biologists," Møller and Jennions state:

---

[55] G. L. Stebbins, "In Defense of Evolution: Tautology or Theory?" *American Naturalist* 111 (1997): 386-390; Ernst Mayr, *The Growth of Biological Thought: Diversity, Evolution, and Inheritance* (Cambridge: Harvard University Press, 1982); J. F. Quinn and A. E. Dunham, "On Hypothesis Testing in Ecology and Evolution," *American Naturalist* 122 (1983): 602–617; L. B. Slobodkin, "Intellectual Problems of Applied Ecology," *BioScience* 38 (1988): 337–342; P. Y. Quenette and J. F. Gerard, "Why Biologists do not Think like Newtonian Physicists," *Oikos* 68 (1993): 361-363; Michael Begon, "The Vole *Clethrionomys rufocanus*—A Modern Classic?" *Researches on Population Ecology* 40 (1998): 145-147; B. Brecking and Q. Dong, "Uncertainty in Ecology and Ecological Modeling," in *Handbook of Ecosystem Theories and Management* (Boca Raton: CRC Press, 2000), 51-73.

[56] D. Lindenmayer and M. Hunter, "Some Guiding Concepts for Conservation Biology," *Conservation Biology* 24 (2010): 1459–1468.

> When first conducting research many graduate students are disappointed when they encounter the fact that biologists explain so little of the variance in their data … Thus the naïve question is as follows: Can we ever explain 100% of the variance? The obvious answer is no, and there are several reasons why that is the case. In particular, biology differs from many other subjects in the natural sciences by being considerably more complex, with consequences for the amount of variation that can be explained by observational or experimental studies.[57]

They then summarized two ideas that are relatively uncontested in the ecological sciences and that help explain the low R^2 values of ecological models. First, the idea that organisms respond to too many variables simultaneously for ecologists to measure. Second, the idea that environments are "random" and "unpredictable."

The first idea, in other words, is that ecological models would better explain natural phenomena *if only* the ecologist had the ability to measure and incorporate more variables. When looking at a messy scatterplot, then, ecologists see phantom variables—variables that were not measured yet undoubtedly (in their view) shape the relationships they seek to understand. Indeed, the discussion sections of ecology articles often contain statements like "we did not measure rates of fish growth, ingestion, or assimilation; variation in these rates could decouple the relationship between diet P content and consumer P excretion rates;" or, "these mechanisms are probably resource heterogeneity and patchiness, though we did not measure these directly in our studies."[58] Such claims pertain to the *number of* variables in the natural world.

The second idea posits that nature is intrinsically complex, and that therefore it is difficult to visualize patterns among variables. As philosopher Elliot Sober has argued, in ecology "variation is not thought of as a deflection from the natural state of uniformity. Rather, variation is taken to be a fundamental property in its own right."[59] In short, ecologists *expect* their graphs to be messy. Indeed, they are not alone in assuming that biological entities are more various than abiotic entities. Historians Gerald Geison and Manfred Laubichler have written, for example, "organisms, or living systems in general, vary to an extent that, say, hydrogen atoms do not."[60] Such claims pertain to the *nature* of variables in the natural world.

Møller and Jennions fail to mention a third source of purported messiness in ecology. Historically, ecologists have also described the events that lead to any ecological assemblage

---

[57] Møller and Jennions, "How Much Variance can be Explained," 493.

[58] G. E. Small, C. M. Pringle, M. Pyron, and J. H. Duff, "Role of the Fish *Astyanax aeneus* (Characidae) as a Keystone Nutrient Recycler in Low-nutrient Neotropical Streams," *Ecology* 92 (2011): 386-397, 392; M. G. St. John, D. H. Wall, and H. W. Hunt, "Are Soil Mite Assemblages Structured by the Identity of Native and Alien Grasses?" *Ecology* 87 (2006): 1314-1324, 1319.

[59] Elliott Sober, "Philosophical Problems for Environmentalism," in *The Preservation of Species* (Princeton: Princeton University Press, 1986), 173-194.

[60] Gerald L. Geison and Manfred D. Laubichler, "The Varied Lives of Organisms: Variation in the Historiography of the Biological Sciences," *Studies in the History and Philosophy of Biological and Biomedical Sciences*, 32 (2001): 1-29, 2.

as "unpredictable," "unique," and "contingent."[61] In 1959, for instance, the zoologist Ernst Mayr wrote, "The more I study evolution the more I am impressed by the uniqueness, by the unpredictability, and by the unrepeatability of events ... Is it not perhaps a basic error of methodology to apply such a generalizing technique as mathematics to a field of unique events?"[62] Forty years later, ecologist John Lawton argued that contingency "makes it difficult, indeed, virtually impossible, to find patterns that are universally true in ecology."[63] Thus, in this third respect, too, ecologists have naturalized organismal and environmental variation. A popular statistical textbook by Gotelli and Ellison explains that unlike laboratory scientists, who often assume that "eliminating measurement error and contamination will lead to clean and repeatable data that are correct," ecologists believe their discipline's models will be messy in perpetuity.[64] Such claims pertain to the *contingency* of variables in the natural world.

Considering these sources and the history of linear regression simultaneously, it becomes clear that not only has ecologists' use of linear regression shaped understandings of nature, but ecologists' understandings of nature have shaped their use of linear regression. Ecologists were inclined to accept models with low R^2 values because they did not expect to find tight correlation between organisms and their environments—both of which were seen as various. In turn, low R^2 values naturalized complexity and messiness.

STS and Statistical Practices

Two bodies of work in STS are particularly relevant to the argument that ecological variation and low R^2 values are mutually shaping—work on the visualization of data, and work on the construction of models. Historians and philosophers of science have persuasively argued that the "mathematization" of natural objects through graphs, tables, and formulas is a crucial aspect of "scientific seeing," and that visualizations privilege particular ways of seeing the world.[65] Historically, scientists have conceptualized models as arbitrators between theory and

---

[61] Møller and Jennions, "How Much Variance can be Explained." Not to mention ecologists' argument that the long time scales, lack of replication, and lack of controls inherent in many ecological studies prevent effective use of the classic reductionist approach. See E. McCoy, "Philosophies of Evidence Encounter the Realities of Data," in *The Nature of Scientific Evidence: Statistical, Philosophical, and Empirical Considerations* (Chicago: University of Chicago Press, 2004), 97-99.

[62] Mayr, *The Growth of Biological Thought*, 317.

[63] J. Lawton, "Are there General Laws in Ecology?" *Oikos*, 84 (1999): 177-192, 179.

[64] Nicholas J. Gotelli and Aaron M. Ellison, *Primer on Ecological Statistics* (New York: Sinauer Associates, 2004), 10. Murray (2000), in turn, argued that what Lawton (1999) referred to as "contingencies" were in fact "initial conditions." The variety of initial conditions, he continues, "does not preclude the existence of universal laws." On the distinction between "contingent" and "contingent upon." B. G. Murray, "Universal Laws and Predictive Theory in Ecology and Evolution," *Oikos* 89 (2000): 403-408. See also J. Smith and C. Jenks, "Complexity, Ecology and the Materiality of Information," *Theory, Culture & Society* 22 (2005): 141-163.

[65] Michael Lynch, "The Externalized Retina: Selection and Mathematization in the Visual Documentation of Objects in the Life Sciences," *Human Studies* 11 (1988): 201-234; Donna Haraway, *Modest_Witness@Second_Millennium.FemaleMan©_Meets_Oncomouse: Feminism and Technoscience* (New York: Routledge, 1997); Luc Pauwels, *Visual Culture of Science: Rethinking Representational Practices in Knowledge Building and Science Communication* (Hanover: Dartmouth

data, a sort of midpoint on the path to visualization.[66] Theories can be validated or rejected in their entirety, whereas models are deemed more or less useful. In short, models are given the status of tools.

Like tools, models occupy an interesting space at the intersection of the material and the conceptual. In constructing models, scientists engage in multiple stages of negotiation.[67] The decision of which variables to include in a model, for example, may be influenced (or determined) by the availability of data, the precedent of prior research, and the opinions of colleagues.[68] But in the case of models, evidence of this negotiation can be difficult to find. Models in published articles, and especially widespread statistical models like linear regression, appear to be resolved things. The negotiation happens in the field, in graduate training, in peer review.

As a field science, ecology presents scholars in STS and related humanities and social science fields with an opportunity to study new aspects of models. An ecological fact—say, "the diameter of lodgepole pines decreases with altitude"—can be traced through multiple

---

College Press, 2005); Catelijne Coopmans, Janet Vertesi , Michael E. Lynch and Steve Woolgar, *Representation in Scientific Practice Revisited* (Cambridge: MIT Press, 2013).

[66] Sergio Sismondo, "Models, Simulations, and Their Objects," *Science in Context* 12 (1999): 247-260.

[67] Mary Hesse, *Models and Analogies in Science* (Notre Dame: Notre Dame Press, 1966); P. Keating, A. Cambrosio, and M. MacKenzie. "The Tools of the Discipline: Standards, Models, and Measures in the Affinity/Avidity Controversy in Immunology," in *The Right Tools for the Job: At Work in Twentieth-Century Life Sciences* (Princeton: Princeton University Press, 1992), 312-356; Joan H. Fujimura, "Crafting Science: Standardized Packages, Boundary Objects and 'Translation,'" in *Science as Practice and Culture* (Chicago: University of Chicago Press, 1992), 168-211; Theodore M. Porter, "Objectivity as Standardization: The Rhetoric of Impersonality in Measurement, Statistics and Cost-benefit Analysis," in *Rethinking Objectivity* (Durham: Duke University Press, 1994), 197-237; Geoffrey C. Bowker and Susan Leigh Star, *Sorting Things Out: Classification and its Consequences* (Cambridge: MIT Press, 1999).

[68] In disciplines such as molecular and cell biology, entire organisms have served as model systems. An expanding subfield of STS literature explores the construction and deployment of biological model species such as flies and mice. See R. White and T. Caskey, "The Human as an Experimental System in Molecular Genetics," *Science* 240 (1988): 1483-1488; B. Kimmelman, "Organisms and Interests in Scientific Research: R.A. Emerson's Claims for the Unique Contributions of Agricultural Genetics," in *The Right Tools for the Job: At Work in Twentieth-Century Life Sciences* (Princeton: Princeton University Press, 1992), 163-204; Gregg Mitman and Anne Fausto-Sterling, "Whatever Happened to *Planaria*? C. M. Child and the Physiology of Inheritance," in *The Right Tools for the Job: At Work in Twentieth-Century Life Sciences* (Princeton: Princeton University Press, 1992), 172-197; Robert E. Kohler, *Lords of the Fly: Drosophila Genetics and the Experimental Life* (Chicago: Chicago University Press, 1994); C. R. Stillwell, "Thymectomy as an Experimental System in Immunology," *Journal of the History of Biology* 27 (1994): 379–401; Angela Creager and Gerald Geison, "Research Materials and Model Organisms in the Biological and Biomedical Sciences," *Studies in History and Philosophy of Biological and Biomedical Sciences* 30 (1999): 315–318; Rachel A. Ankeny, "Model Organisms as Models: Understanding the 'Lingua Franca' of the Human Genome Project," *Proceedings of the Philosophy of Science Association* 68 (2001): S251-S261; I. Löwy and J. Gaudillière, "Disciplining Cancer: Mice and the Practice of Genetic Purity," in *The Invisible Industrialist* (New York: Macmillan, 1998), 209-249; Angela Creager, *The Life of a Virus: TMV as an Experimental Model, 1930-1965* (Chicago: University of Chicago Press, 2001); Karen Rader, *Making Mice: Standardizing Animals for American Biomedical Research, 1900-1955* (Princeton: Princeton University Press, 2004).

stages of negotiation. Imagine a graduate student (1) deciding where she would conduct her research, (2) constructing a sampling design, and deciding which variables to measure, (3) kneeling in the mud with a tape measure, adjusting it around the trunk of a lodgepole pine, recording data on rain-proof paper, (4) choosing which data to model, (5) entering that data into a statistical model that she has chosen to employ, a linear regression, perhaps, (6) interpreting both the visual and numerical output of the linear regression model, and, despite the low R^2 value of the graph with "altitude" on the x-axis and "diameter" on the y-axis, concluding the graph displays a relationship between the two variables.

Previous scholarship on model construction in ecology and other disciplines has tended to consider the decision to model, or the choice of variables to include in a model, rather than the interpretation of model results.[69] Martin *et al*. analyzed what would correspond to stage (1) above: they found that most ecological fieldwork is conducted in protected areas.[70] Pertaining to stage (3), Roth and Bowen described the processes through which ecologists used tags, tables, and maps to transform their observations of desert lizards into rows and columns on spreadsheets.[71] Klingle explored stage (4), reconstructing the process through which ecologists decided which variables to include and which to exclude in models of the "complex" Fern Lake ecosystem in Washington.[72] This article has examined aspects of stage (5) —the historical contingencies that led ecologists to practice linear regression—and stage (6) — the details of how ecologists have interpreted the output of linear regression.

In employing linear regression, ecologists had to agree not only on which variables to model, but also on the nature of the relationships *among* those variables. Confronted with the environmental disaster of the Dust Bowl, for example, American ecologists expected to develop useful explanations without the confidence that those explanations would enable them fully to predict or avert similar future events. They expected relationships among organismal and environmental variables to be difficult to perceive. Their use of linear regression both depended on and reproduced the idea that nature is complex, unruly, entangled, yet knowable.

---

[69] Relatedly, Levin (2014) explored how contemporary metabolics researchers have developed multifactorial understandings of metabolism through multivariate statistical analyses. The notion of metabolism as a complex process is not waiting to be discovered, she argued, but instead is actively created and enacted by scientists. Nadine Levin, "Multivariate Statistics and the Enactment of Metabolic Complexity," *Social Studies of Science* 44 (2014): 555-578.

[70] Laura J. Martin, Bernd Blossey, and Erle C. Ellis, "Mapping Where Ecologists Work: Biases in the Global Distribution of Terrestrial Ecological Observations," *Frontiers in Ecology and the Environment* 10 (2012): 195-201.

[71] W. Roth and G. M. Bowen, "Digitizing Lizards: The Topology of 'Vision' in Ecological Fieldwork," *Social Studies of Science* 29 (1999): 719-764; John Law and Michael Lynch, "Lists, Field Guides, and the Descriptive Organization of Seeing: Birdwatching as an Exemplary Observation Activity," in *Representation in Scientific Practice* (Cambridge: MIT Press, 1990).

[72] Matthew Klingle, "Plying Atomic Waters: Lauren Donaldson and the "Fern Lake Concept" of Fisheries Management," *Journal of the History of Biology* 31 (1998): 1–32.

Conclusion

Since the publication of Darwin's *Origin*, ecologists have sought patterns in complexity. Linear regression has been an important part of this history. By 1990, linear regression was arguably the most widely used statistical method in ecology.[73] Clearly, linear regression has been important to the production of ecological knowledge. And given the influence of ecology on environmental policy worldwide, it is important to understand how ecological knowledge has been and is being produced.

What is at stake in characterizing the natural world as complex yet interrelated? A lot, in a world where governments often justify their actions on the basis of "best available science." Ecologists frequently testify before Congress, in courts, and in the media.[74] As spokespeople for nature,[75] ecologists claim authoritative knowledge not only about how the natural world *is* structured, but also how it *ought to be* structured. Ecological theory has thus come to shape conservation and environmental policy. The choice to read biological and environmental variables as complex rather than simple, and interdependent rather than independent, therefore has clear political stakes.[76] By characterizing the natural world as complex, ecologists have successfully argued for the exclusivity and necessity of their discipline. (Recall Taylor in 1935: "Who but a geo-bio-ecologist, one who knows something of interrelationships and of plant and animal indicators and soils, is qualified for the important tasks of land classification?")

The argument that ecological variables are correlated, meanwhile, became central to environmental regulation and legislation, including the banning of DDT in 1972 and the Endangered Species Act of 1973. Ecologist Orie Loucks' 1972 guide to delivering expert testimony, for example, emphasized the necessity of convincing the jury of "the quality of interconnections that couple air, land, and water systems, and man's long-range impacts on them."[77] Into the 1980s, ecological knowledge increasingly informed environmental regulation and biodiversity conservation. Much of this knowledge came from linear regression. For example, in the late 1980s ecologists used linear regression to argue that certain private lands

---

[73] J. Wilson White, Andrew Rassweiler, Jameal F. Samhouri, Adrian C. Stier, Crow White, "Ecologists Should Not use Statistical Significance Tests to Interpret Simulation Model Results," *Oikos* 123 (2014): 385-388.

[74] Arthur F. McEvoy, *The Fisherman's Problem: Ecology and Law in the California Fisheries, 1850-1980* (Cambridge: Cambridge University Press, 1986); Stephen Bocking, *Nature's Experts: Science, Politics, and the Environment* (New Brunswick: Rutgers University Press, 2004).

[75] Bruno Latour, *Politics of Nature: How to Bring the Sciences into Democracy* (Cambridge: Harvard University Press, 2004).

[76] Evans (1999) and van den Bogaard (1999) have shown how economic models concretize assumptions that have moral weight. Robert Evans, "Economic Models and Policy Advice: Theory Choice or Moral Choice?" *Science in Context* 12 (1999): 351-376; Adrienne van den Bogaard, "The Cultural Origins of the Dutch Economic Modeling Practice," *Science in Context* 12 (1999): 333-350.

[77] Orie L. Loucks, "Systems Methods in Environmental Court Actions," in *Systems Analysis and Simulation in Ecology Vol. II* (New York: Academic Press, 1972), 419-472, 424. Also V. J. Yannacone, "Plaintiffs' Brief in the Project Rulison Case," *Cornell Law Review* 55 (1974): 761-807. On the intersections of science and law, see Sheila Jasanoff, *Science at the Bar: Law, Science, and Technology in America* (Cambridge: Harvard University Press, 1995).

in central Florida were suitable for re-colonization by the endangered Florida panther. Interpretations of their results were debated widely.[78]

Some historians of science have argued that by embracing statistical methods, ecologists abandoned their intellectual tradition. Donald Worster maintained that in the 1960s, "For the first time, mathematicians could see in ecology the opportunity to quantify."[79] Paolo Pallandino argued that "from the 1920s onward, ecologists have sought to transform the older traditions of natural history and natural resource management into a rigorous scientific discipline, and have done so by staking the legitimacy of their endeavor more or less explicitly on approximating the approach of physicists."[80] Michael Barbour described ecologists as "actors in the long-running story of holism yielding to reductionism, a theme in the history of science."[81] Sharon Kingsland wrote that "ecologists continue to look toward mathematics and the physical sciences for ideas, techniques, and models of what science should be."[82]

Alternatively, I have argued that although ecologists gradually incorporated linear regression into their practices from 1930 to 1960, they maintained their disciplinary commitment to the idea of nature's complexity. This commitment is inscribed in how ecologists have employed linear regression. For many ecologists, a low R^2 value was not a reason to dismiss a model, but rather, proof that the model reflected the natural world.[83] Indeed, graphs with high R^2 values are suspect in ecology, dismissed as fraudulent or as artifacts of "over-fitting" (including too many variables in a model, which artificially improves the fit of the model).[84] As Thomas Gieryn noted, emulation of the laboratory sciences is not the only means

---

[78] K. S. Shrader-Frechette and E. D. McCoy, *Method in Ecology: Strategies for Conservation* (Cambridge: Cambridge University Press, 1992).

[79] Donald Worster, *Nature's Economy: A History of Ecological Ideas* (Cambridge: Cambridge University Press, 1977), 304.

[80] Paolo Palladino, "Defining Ecology: Ecological Theories, Mathematical Models, and Applied Biology in the 1960s and 1970s," *Journal of the History of Biology* 24 (1991): 223-243.

[81] Michael G. Barbour, "Ecological Fragmentation in the Fifties," in *Uncommon Ground: Rethinking the Human Place in Nature* (New York: W. W. Norton, 1996), 233-255.

[82] Sharon Kingsland, *Modeling Nature: Episodes in the History of Population Ecology* (Chicago: University of Chicago Press, 1995), 234.

[83] In *Politics of Nature*, Latour has powerfully argued that ecology "allowed us to dispense with the requirements of discussion and due process in building the common world" and has cast nature as "a hidden procedure for apportioning speech and authority." What is often at stake in ecological arguments, Williams (1980) suggests, is "the ideas of different kinds of societies." One question, then, is whether "messy" implies "uncontrollable." It would be interesting to read gender into this: The feminized natural world is described as messy, chaotic, and multivariate, yet ultimately constrained by deterministic (masculinized) laws. R. Williams, *Problems in Materialism and Culture* (London: Verso, 1980).

[84] Overfitting is defined as the phenomenon of a statistical model describing random error instead of the underlying relationship, usually because the modeler has included too many parameters relative to the number of observations. Statisticians, and increasingly ecologists, use methods like cross-validation, Bayesian priors, and model comparison to attempt to balance precision and parsimony. These techniques either penalize overly complex models or test the model's ability to generalize by evaluating its performance on a set of previously unseen data. While a deep analysis of these techniques is beyond the scope of this article, it would be interesting to compare these narratives about

through which field sciences have acquired credibility. In some disciplines, the field carries with it an idea of unadulterated reality, so that "an inevitable lack of control becomes its own virtue." [85] Today "ecological" is practically synonymous with "messy," "contingent," "entangled."

In analyzing the messiness of ecologist's graphs, I do not imply that ecology is less rigorous than other sciences. Many disciplines value messy graphs; economics comes to mind. Nor is ecology the only discipline to employ linear regression. Helene Wagner has noted that "linear regression is the workhorse of statistical modeling in many disciplines, including such disparate fields as ecology, social sciences, or econometrics and finance."[86] Rather than use low R^2 values to categorize these fields as "soft sciences," one can use them to interrogate the soft, medium, and hard norms employed in any scientific discipline and to question why certain norms are flexible while others are recalcitrant. Difference in the interpretation of seemingly standardized methods such as linear regression can illuminate the construction and deployment of scientific knowledge.

Although linear regression is used across the natural and physical sciences, it would be a mistake to treat it as a stable practice. In incorporating linear regression into their practices, ecologists applied it to new types of questions, relating organismal variables to environmental variables. And in interpreting linear regression, ecologists saw pattern in spite of messiness. With polygons of variation, variation had been the thing on display and the object to be analyzed. With linear regression, however, correlation became the object to be analyzed and variation became background noise. In this way, variation was thoroughly naturalized.[87]

Ecology's low R^2 values offer an entry point to exploring the production of ecological knowledge. In the mid-20th century, ecologists gradually assented to the importance of statistical practices and mathematization. But they did not abandon complexity for simplification. Instead, they employed linear regression in accordance with their tacit beliefs.

"error" and "complexity" with those around R^2 values. In itself, the idea of balancing precision and parsimony is related to my analysis of constrained messiness.

[85] Thomas F. Gieryn, "City as Truth-spot Laboratories and Field-sites in Urban Studies," *Social Studies of Science* 36 (2006): 5-38. Today many ecologists believe that field studies produce "truer" results than laboratory studies. Ecologist Stephen Carpenter (1996), for instance, argued that meaningful ecological work stems from "deep appreciation of natural history and real ecosystems, which can come from extensive field experience but not from the campus." Stephen R Carpenter, "Microcosm Experiments have Limited Relevance for Community and Ecosystem Ecology," *Ecology* 77 (1996): 677-680.

[86] Helene Wagner, "Rethinking the Linear Regression Model for Spatial Ecological Data," *Ecology* 94 (2013): 2381-2391.

[87] It should be noted that quite recently, ecologists in the subfield of plant-insect interactions have begun to promote the study of "interspecific variation." They frame it as a new research field, though it has as predecessors the 19th century practices that I have discussed. Take, for example, studies of intraspecific variation in plant chemical defenses (Agrawal *et al*., 2012), the move to incorporate noise into population models (Vasseur and Yodzis, 2004), or, mind-bendingly, efforts to find generalizable laws of natural complexity (Bak and Paczuski, 1995). Anurag A. Agrawal, Amy Hastings, Marc Johnson, John Maron, and Juha-Pekka Salminen, "Insect Herbivores Drive Real-time Ecological and Evolutionary Change in Plant Populations," *Science* 338 (2013): 113-116; D. A. Vasseur and P. Yodzis, "The Color of Environmental Noise," *Ecology* 85 (2004): 1146-1152; Per Bak and Maya Paczuski, "Complexity, Contingency, and Criticality," *Proceedings of the National Academy of Sciences USA* 92 (1995): 6689-6696.

In a purportedly complex and interdependent world, ecologists assumed that the variance accounted for by any single variable *should be* small. Variation was taken to be an inherent characteristic of the object of study. In this sense, ecological studies employing linear regression or related statistical methods —in other words, most ecological studies—are neither "holistic" nor simply "reductionist." Messy graphs make simultaneous claims to generalization and to specificity.[88] Ecologists' willingness to see relationships in messy graphs speaks to a historical and contemporary tension in ecology between the universal and the specific, the simple and the complex, the determined and the contingent—the simultaneous embrace of Darwin's "entangled bank" and the "laws that act around us."

Laura Jane Martin is an environmental historian and ecologist. She is currently a Ziff Environmental Fellow at the Harvard University Center for the Environment and a postdoctoral fellow in Harvard's Department of the History of Science. Her current project, *Designing Nature: Ecological Restoration from the Dust Bowl to De-extinction*, explores the emergence of restoration ecology as an idea, practice, and scientific discipline in the 20th century United States.

ACKNOWLEDGEMENTS This article benefited from the helpful feedback of Sara Pritchard, Suman Seth, Clifford Kraft, Aaron Sachs, Paul Nadasdy, Alex Alexiades, Ezra Feldman, Matthew Chrulew, Thom van Dooren, the Cornell Science Sciences Research Group, and three anonymous reviewers. LJM was supported by the National Science Foundation Grant No. 1329750 and by the Cornell University Society for the Humanities.

Bibliography

Agrawal, Anurag A., Amy Hastings, Marc Johnson, John Maron, and Juha-Pekka Salminen. "Insect Herbivores Drive Real-time Ecological and Evolutionary Change in Plant Populations." *Science* 338 (2012): 113-116.

Ankeny, Rachel A. "Model Organisms as Models: Understanding the 'Lingua Franca' of the Human Genome Project." *Proceedings of the Philosophy of Science Association* 68 (2001): S251-S261.

Anker, Peder. *Imperial Ecology: Environmental Order in the British Empire, 1895-1945*. Cambridge: Harvard University Press, 2002.

Archdeacon, Thomas J. *Correlation and Regression Analysis: A Historian's Guide*. Madison: University of Wisconsin Press, 1994.

Bak, Per, and Maya Paczuski. "Complexity, Contingency, and Criticality." *Proceedings of the National Academy of Sciences USA* 92 (1995): 6689-6696.

[88] The idea that variation thwarts generalization, even though variation itself is supposedly a result of law (natural selection), remains the unresolved (and underexplored) core of ecological science. For an excellent essay on the ideas of chaos and order in ecology, see Christopher H. Eliot, "The Legend of Order and Chaos: Communities and Early Community Ecology," *Handbook of the Philosophy of Science* 11 (2011): 49-107.

Barbour, Michael G. "Ecological Fragmentation in the Fifties." In *Uncommon Ground: Rethinking the Human Place in Nature*, edited by William Cronon, 233-255. New York: W. W. Norton, 1996.

Begon, Michael. "The Vole *Clethrionomys rufocanus*—A Modern Classic?" *Researches on Population Ecology* 40 (1998): 145-147.

Blyth, Stephen. "Karl Pearson and the Correlation Curve." *International Statistical Review* 62 (1994): 393-403.

Bocking, Stephen. *Nature's Experts: Science, Politics, and the Environment*. New Brunswick: Rutgers University Press, 2004.

Bowker, Geoffrey C., and Susan Leigh Star. *Sorting Things Out: Classification and its Consequences.* Cambridge: MIT Press, 1999.

Brecking, B., and Q. Dong. "Uncertainty in Ecology and Ecological Modeling." In *Handbook of Ecosystem Theories and Management*, edited by S. E. Jorgensen and Muller Felix, 51-74. Boca Raton: CRC Press, 2000.

Browne, Janet. *The Secular Ark: Studies in the History of Biogeography.* New Haven: Yale University Press, 1983.

Burgess, Robert. *Historical Data and Some Preliminary Analyses*. Washington D.C.: Ecological Society of America, c. 1976.

Callon, Michel. 1986. "Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of St Brieuc Bay." In *Power, Action and Belief: A New Sociology of Knowledge?* edited by John Law, 196-233. London: Routledge & Kegan Paul, 1986.

Camic, Charles C., and Yu Xie. "The Statistical Turn in American Social Science: Columbia University, 1890 to 1915." *American Sociological Review* 59 (1994): 773-805.

Carpenter, Stephen R. "Microcosm Experiments have Limited Relevance for Community and Ecosystem Ecology." *Ecology* 77 (1996): 677-680.

Cittadino, Eugene. *Nature as the Laboratory: Darwinian Plant Ecology in the German Empire, 1880-1900.* Cambridge: Cambridge University Press, 1990.

Clements, Frederic E. "Experimental Ecology in the Public Service." *Ecology* 16 (1935): 342-363.

Cook, C. Wayne. "The Use of Multiple Regression and Correlation in Biological Investigations." *Ecology* 41 (1960): 556-560.

Cornwell, Myriah L., and Lisa M. Campbell. "Co-producing Conservation and Knowledge: Citizen-based Sea Turtle Monitoring in North Carolina, USA." *Social Studies of Science* 42 (2012): 101-120.

Cox, D. R. "Biometrika: The First 100 Years." *Biometrika* 88 (2011): 3-11.

Creager, Angela, and Gerald Geison. "Research Materials and Model Organisms in the Biological and Biomedical Sciences." *Studies in History and Philosophy of Biological and Biomedical Sciences* 30 (1999): 315–318.

Creager, Angela. *The Life of a Virus: TMV as an Experimental Model, 1930-1965*. Chicago: University of Chicago Press, 2001.

Csada, R. D., P. C. James, and R. Espie. "The "File Drawer Problem" of Non-significant Results: Does it Apply to Biological Research?" *Oikos* 76 (1996): 591–593.

Darwin, Charles. *On the Origin of Species.* London: John Murray, 1859.

Diller, Oliver D. "The Relation of Temperature and Precipitation to the Growth of Beech in Northern Indiana," *Ecology* 16 (1935): 72-81.

Draper, Norman R., and H. Smith. *Applied Regression Analysis*. New York: John Wiley and Sons Inc., 1966.

Easterbrook, P. J., R. Goplan, J. A. Berlin, and D. R. Matthews. "Publication Bias in Clinical Research." *The Lancet* 337 (1991): 867-872.

Ecological Society of America. "Bibliography of Quantitative Work Bearing on Ecological Problems." *Ecology* 22 (194): 343-344.

Edgerton, Frank N. ed. *History of American Ecology.* New York: Arno Press, 1977.

Editors. "Editorial: The Spirit of Biometrika," *Biometrika* 1 (1901): 3-6.

Edmondson, W. T. "Reproductive Rate of Planktonic Rotifers as Related to Food and Temperature in Nature." *Ecological Monographs* 35 (1965): 61-111.

Fanelli, Daniele. "Negative Results are Disappearing from Most Disciplines and Countries." *Scientometrics* 90 (2012): 891-904.

Ferrill, W. K. "Effect of Environmental Conditions on Survival and Growth of Forest Tree Seedlings under Field Conditions in the Piedmont Region of North Carolina." *Ecology* 34 (1953): 667-688.

Fidler, F., G. Cumming, M. Burgman, and N. Thomason. "Statistical Reform in Medicine, Psychology, and Ecology." *The Journal of Socio-Economics* 33 (2004): 615-630.

Field, Alexander J. *A Great Leap Forward: 1930s Depression and U.S. Economic Growth*. New Haven: Yale University Press, 2011.

Filho, F., J. A. Silva, and E. Rocha. "What is R^2 All About?" *Leviathan* 3 (2011): 60-68.

Friendly, M., D. Denis. "The Early Origins and Development of the Scatterplot." *Journal of the History of the Behavioral Sciences* 41 (2005): 103-130.

Fujimura, Joan H. Crafting Science: Standardized Packages, Boundary Objects and 'Translation.' In *Science as Practice and Culture*, edited by Andrew Pickering, 158-211. Chicago: University of Chicago Press, 1992.

Galison, Peter. *Image and Logic: A Material Culture of Microphysics*. Chicago: University of Chicago Press, 1997.

Galton, Francis. "Biometry." *Biometrika* 1 (1901): 7-10.

_____ . "Co-relations and their Measurement, Chiefly from Anthropometric Data." *Proceedings of the Royal Society of London* 45 (1888): 135-145.

_____ . *Natural Inheritance*. New York: Macmillan and Company, 1894.

Geison, Gerald L., and Manfred D. Laubichler. "The Varied Lives of Organisms: Variation in the Historiography of the Biological Sciences." *Studies in the History and Philosophy of Biological and Biomedical Sciences* 32 (2001): 1-29.

Gieryn, Thomas F. "City as Truth-spot Laboratories and Field-sites in Urban Studies." *Social Studies of Science* 36 (2006): 5-38.

Golley, Frank B. *A History of the Ecosystem Concept*. New Haven: Yale University Press, 1993.

Gossel, P. P. "The Need for Standard Methods: The Case of American Bacteriology." In *The Right Tools for the Job: At Work in Twentieth-Century Life Sciences*, edited by Adele Clarke and Joan H. Fujimura, 287-311. Princeton: Princeton University Press, 1992.

Gotelli, Nicholas J., and Aaron M. Ellison. *Primer on Ecological Statistics*. New York: Sinauer Associates, 2004.

Gregory, James. *American Exodus: The Dust Bowl Migration and Okie Culture in California*. New York: Oxford University Press, 1989.

Hacking, Ian. *The Taming of Chance*. Cambridge: Cambridge University Press, 1990.

Hald, Anders. *A History of Parametric Statistical Inference from Bernoulli to Fischer, 1713-1935*. New York: Springer Science, 2007.

Hawley, F. M. "Relationship of Southern Cedar Growth to Precipitation and Run Off." *Ecology* 18 (1937): 398-405.

Hesse, Mary. *Models and Analogies in Science*. Notre Dame: Notre Dame Press, 1966.

Hogle, L. "Standardization across Non-standard Domains: The Case of Organ Procurement." *Science, Technology, and Human Values* 20 (1995): 482-501.

Jordan, Kathleen, and Michael Lynch. "The Sociology of Genetic Engineering Technique: Ritual and Rationality in the Performance of the 'Plasmid Prep.'" In *The Right Tools for the Job: At Work in Twentieth-Century Life Sciences*, edited by Adele Clarke and Joan H. Fujimura, 77-114. Princeton: Princeton University Press.

Keating, P., A. Cambrosio, and M. MacKenzie. "The Tools of the Discipline: Standards, Models, and Measures in the Affinity/Avidity Controversy in Immunology." In *The Right Tools for the Job: At Work in Twentieth-Century Life Sciences*, edited by Adele Clarke and Joan H. Fujimura, 312-356. Princeton: Princeton University Press, 1992.

Kimmelman, B. "Organisms and Interests in Scientific Research: R.A. Emerson's Claims for the Unique Contributions of Agricultural Genetics." In *The Right Tools for the Job: At Work in Twentieth-Century Life Sciences*, edited by Adele Clarke and Joan H. Fujimura, 163-204. Princeton: Princeton University Press, 1992.

Kinchy, Abby J. and Daniel L. Kleinman. "Organizing Credibility: Discursive and Organizational Orthodoxy on the Borders of Ecology and Politics." *Social Studies of Science* 33 (2003): 869-896.

Kingsland, Sharon. *Modeling Nature: Episodes in the History of Population Ecology*. Chicago: University of Chicago Press, 1995.

_____ . *The Evolution of American Ecology, 1890-2000*. Baltimore: Johns Hopkins University Press, 2005.

Klenk, Nicole L., Gordon M. Hickey, and J. I. MacLellan. "Evaluating the Social Capital Accrued in Large Research Networks: The Case of the Sustainable Forest Management Network (1995-2000)." *Social Studies of Science* 40 (2010): 931-960.

Klingle, Matthew. "Plying Atomic Waters: Lauren Donaldson and the "Fern Lake Concept" of Fisheries Management." *Journal of the History of Biology* 31 (1998): 1–32.

Kohler, Robert E. *Landscapes and Labscapes: Exploring the Lab-Field Border in Biology*. Chicago: University of Chicago Press, 2002.

_____ . *Lords of the Fly: Drosophila Genetics and the Experimental Life*. Chicago: Chicago University Press, 1994.

Kwa, Chunglin. "Local Ecologies and Global Science Discourses and Strategies of the International Geosphere-Biosphere Programme." *Social Studies of Science* 35 (2005): 923-950.

Lackey, R. T. "Axioms of Ecological Policy." *Fisheries* 31 (2006): 286-290.

Latour, Bruno. "Where are the Missing Masses? The Sociology of a few Mundane Artifacts." In *Shaping Technology/Building Society: Studies in Sociotechnical Change*, edited by W. Bijker and John Law, 225-258. Cambridge: MIT Press, 1992.

_____ . *Politics of Nature: How to Bring the Sciences into Democracy*. Cambridge: Harvard University Press, 2004.

Law, John, and Michael Lynch. "Lists, Field Guides, and the Descriptive Organization of Seeing: Birdwatching as an Exemplary Observation Activity." In *Representation in Scientific Practice*, edited by Michael Lynch and Steve Woolgar, 267-299. Cambridge: MIT Press, 1990.

Lawton, J. "Are there General Laws in Ecology?" *Oikos* 84 (1999): 177-192.

Lindenmayer, D., and M. Hunter. "Some Guiding Concepts for Conservation Biology." *Conservation Biology* 24 (2010): 1459–1468.

Low-Decarie, Etienne, Corey Chivers, and Monica Granados, "Rising Complexity and Falling Explanatory Power in Ecology." *Frontiers in Ecology and the Environment* 12 (2014): 412-418.

Löwy, I., and J. Gaudillière. "Disciplining Cancer: Mice and the Practice of Genetic Purity." In *The Invisible Industrialist*, edited by I. Löwy and J. Gaudillière, 209-249. New York: Macmillan, 1998.

Lynch, Michael. "The Externalized Retina: Selection and Mathematization in the Visual Documentation of Objects in the Life Sciences." *Human Studies* 11 (1988): 201-234.

MacArthur, Robert H. *Geographical Ecology: Patterns in the Distribution of Species.* New York: Harper & Row, 1972.

Macfayden, Ayman. *Animal Ecology: Aims and Methods*. London: Pittman, 1963.

Maher, Neil. *Nature's New Deal: The Civilian Conservation Corps and the Roots of the American Environmental Movement*. New York: Oxford University Press, 2008.

Mayr, Ernst. *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*. Cambridge: Harvard University Press, 1982.

McCoy, E. "Philosophies of Evidence Encounter the Realities of Data." In *The Nature of Scientific Evidence: Statistical, Philosophical, and Empirical Considerations*, edited by M. Taper and S. Lele, 97-99. Chicago: University of Chicago Press, 2004.

McDonald, J.H. *Handbook of Biological Statistics*. Baltimore: Sparky House Publishing, 2014.

McEvoy, Arthur F. *The Fisherman's Problem: Ecology and Law in the California Fisheries, 1850-1980*. Cambridge: Cambridge University Press, 1986.

Mead, R. "A Note on the Use and Misuse of Regression Models in Ecology." *Journal of Ecology* 59 (1971): 215-219.

Michael, E. L. "Marine Ecology and the Coefficient of Association: A Plea in Behalf of Quantitative Biology." *Journal of Ecology* 8 (1920): 54-59.

Millerand, F., David Ribes, K. S. Baker, and G. C. Bowker. "Making an Issue out of a Standard: Storytelling Practices in a Scientific Community." *Science, Technology, & Human Values* 38 (2013): 7-43.

Mitman, Gregg, and Anne Fausto-Sterling. "Whatever Happened to *Planaria*? C. M. Child and the Physiology of Inheritance." In *The Right Tools for the Job: At Work in Twentieth-Century Life Sciences*, edited by Adele Clarke and Joan H. Fujimura, 172-197. Princeton: Princeton University Press, 1992.

Mitman, Gregg. *The State of Nature: Ecology, Community, and American Social Thought, 1900-1950*. Chicago: University of Chicago Press, 1992.

Moller, A. P., and M. D. Jennions. "How Much Variance can be Explained by Ecologists and Evolutionary Biologists?" *Oecologia* 132 (2002): 492-500.

Moore, Barrington. "The Scope of Ecology." *Ecology* 1 (1920): 3-5.

Murray, B. G. "Universal Laws and Predictive Theory in Ecology and Evolution." *Oikos* 89 (2000): 403-408.

Nassauer, J. I. "Messy Ecosystems, Orderly Frames." *Landscape Journal* 14 (1995): 161-170.

Nelkin, D. "Scientists and Professional Responsibility: The Experience of American Ecologists." *Social Studies of Science* 7 (1977): 75-95.

Nicholls, A. O., and D. M. Calder. "Comments on the Use of Regression Analysis for the Study of Plant Growth." *New Phytologist* 72 (1973): 571-581.

Nyhart, Lynn K. *Biology Takes Form: Animal Morphology and the German Universities, 1800-1900.* Chicago: University of Chicago Press, 1995.

Pahl-Wostl, Claudia. *The Dynamic Nature of Ecosystems: Chaos and Order Entwined*. New York: Wiley, 1995.

Palladino, Paolo. "Defining Ecology: Ecological Theories, Mathematical Models, and Applied Biology in the 1960s and 1970s." *Journal of the History of Biology* 24 (1991): 223-243.

Pauwels, Luc, ed. *Visual Culture of Science: Rethinking Representational Practices in Knowledge Building and Science Communication*. Hanover: Dartmouth College Press, 2005.

Pearson, E. S. *Mathematical Statistics and Data Analysis*. Belmont: Duxbury, 1938.

Pearson, Karl. "Mathematical Contributions to the Theory of Evolution III: Regression, Heredity and Panmixia." *Philosophical Transactions of the Royal Society of London* 187 (1896): 253-318.

_____ . "Notes on the History of Correlation." *Biometrika* 13 (1920): 25-45.

_____ . *Francis Galton: A Centenary Appreciation*. Cambridge: Cambridge University Press, 1922.

Pickering, Andrew. *The Mangle of Practice: Time, Agency, and Science*. Chicago: University of Chicago Press, 1995.

Playfair, William. *The Commercial and Political Atlas*. London, 1786.

Porter, Theodore M. "Objectivity as Standardization: The Rhetoric of Impersonality in Measurement, Statistics and Cost-benefit Analysis." In *Rethinking Objectivity,* edited by A. Megill, 197-237. Durham: Duke University Press, 1994.

_____ . *The Rise of Statistical Thinking, 1820-1900.* Princeton: Princeton University Press, 1986.

_____ . *Trust In Numbers: The Pursuit of Objectivity in Science and Public Life.* Princeton: Princeton University Press, 1995.

Quenette, P. Y., J. F. Gerard JF. "Why Biologists do not Think like Newtonian Physicists." *Oikos* 68 (1993): 361-363.

Quinn, J. F., and A. E. Dunham. "On Hypothesis Testing in Ecology and Evolution." *American Naturalist* 122 (1983): 602–617.

Roney, H. B. "The Effect of Temperature and Light on Oxygen Consumption and Rate of Development of *Helisoma*." *Ecology* 24 (1943): 218-243.

Rorty, R. *Philosophy and the Mirror of Nature*. Princeton: Princeton University Press, 1979.

Rosenthal, R. "The File Drawer Problem and Tolerance for Null Results." *Psychological Bulletin* 86 (1979): 638-641.

Roth, W., and G. M. Bowen. "Digitizing Lizards: The Topology of 'Vision' in Ecological Fieldwork." *Social Studies of Science* 29 (1999): 719-764.

Sachs, Aaron. *The Humboldt Current: Nineteenth Century Exploration and the Roots of American Environmentalism.* New York: Viking, 2006.

Scandura, Jani. *Down in the Dumps: Place, Modernity, American Depression.* Duke University Press, 2008.

Schneider, D. W. "Local Knowledge, Environmental Politics, and the Founding of Ecology in the United States: Stephen Forbes and 'The Lake as Microcosm' (1887)." *Isis* 91 (2000), 681-705.

Scott, James C. *Seeing like a State: How Certain Schemes to Improve the Human Condition Have Failed.* New Haven: Yale University Press, 1999.

Shapin Steven, and Simon Schaffer. *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life.* Princeton: Princeton University Press, 1985.

Shelford, Victor E. "The Organization of the Ecological Society of America 1914-19." *Ecology* 19 (1938): 164-166.

Shrader-Frechette, Kristin, and Earl D. McCoy. "How the Tail Wags the Dog: How Value Judgments determine Ecological Science." *Environmental Values* 3 (1994): 107-120.

Slobodkin, L. B. "Intellectual Problems of Applied Ecology." *BioScience* 38 (1988): 337–342.

Small, G. E., C. M. Pringle, M. Pyron, J. H. Duff. "Role of the Fish *Astyanax aeneus* (Characidae) as a Keystone Nutrient Recycler in Low-nutrient Neotropical Streams." *Ecology* 92 (2011): 386-397.

Smith, J., and C. Jenks. "Complexity, Ecology and the Materiality of Information." *Theory, Culture & Society* 22 (2005): 141-163.

Snedecor, George Waddel. *Statistical Methods Applied to Experiments in Agriculture and Biology* Ames: The Iowa State College Press, 1940.

Sober, Elliott. "Philosophical Problems for Environmentalism." In *The Preservation of Species*, edited by B. Norton, 173-194. Princeton: Princeton University Press, 1986.

St. John, M. G., D. H. Wall, and H. W. Hunt. "Are Soil Mite Assemblages Structured by the Identity of Native and Alien Grasses?" *Ecology* 87 (2006): 1314-1324.

Star, Susan Leigh, and James R. Griesemer. "Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39." *Social Studies of Science* 19 (1989): 387-420.

Stauffer, R. C. "Haeckel, Darwin, and Ecology." *Quarterly Review of Biology* 32 (1957): 138-414.

Stebbins, G. L. "In Defense of Evolution: Tautology or Theory?" *American Naturalist* 111 (1977): 386-390.

Sterling, T. D., W. L. Rosenbaum, and J. J. Weinkam. "Publication Decisions Revisited—The Effect of the Outcome of Statistical Tests on the Decision to Publish and Vice Versa." *American Statistician* 49 (1995): 108-112.

Stigler, S. M. "Francis Galton's Account of the Invention of Correlation." *Statistical Science* 4 (1989): 73-79.

_____ . *The History of Statistics: The Measurement of Uncertainty Before 1900*. Cambridge: Harvard University Press, 1986.

Stillwell, C. R. "Thymectomy as an Experimental System in Immunology." *Journal of the History of Biology* 27 (1994): 379–401.

Tansley, Arthur G. "The Early History of Modern Plant Ecology in Britain," *Journal of Ecology* 35 (1947): 130-137.

Taylor, Norman. "The Beginnings of Ecology," *Ecology* 19 (1938): 352.

Taylor, W. P. "What is Ecology and What Good is it?" *Ecology* 17 (1936): 333-346.

Tobey, Ronald C. *Saving the Prairies: The Life Cycle of the Founding School of American Plant Ecology*. Berkeley: University of California Press, 1981.

Trombulak, S. C., T. L. Fleischner. "Natural History Renaissance." *Journal of Natural History Education and Experience* 1 (2007): 1-4.

Vasseur, D. A., and P. Yodzis. "The Color of Environmental Noise." *Ecology* 85 (2004): 1146-1152.

Vetter, Jeremy. "Introduction." In *Knowing Global Environments: New Historical Perspectives on the Field Sciences*, edited by Jeremy Vetter, 1-16. New Brunswick: Rutgers University Press, 2011.

Wagner, H. H. "Rethinking the Linear Regression Model for Spatial Ecological Data." *Ecology* 94 (2013): 2381-2391.

Weldon, R. "The Variations Occurring in Certain Decapod Crustacea I. *Cragon vulgaris*." *Proceedings of the Royal Society* 47 (1889): 286-291.

White, J. Wilson, Andrew Rassweiler, Jameal F. Samhouri, Adrian C. Stier, and Crow White, "Ecologists Should Not use Statistical Significance Tests to Interpret Simulation Model Results," *Oikos* 123 (2014): 385-388.

White, R., and T. Caskey. "The Human as an Experimental System in Molecular Genetics." *Science* 240 (1988): 1483-1488.

Williams, E. J. *Regression Analysis*. New York: John Wiley and Sons Inc., 1959.

Williams, R. *Problems in Materialism and Culture*. London: Verso, 1980.

Worster, Donald. *Dust Bowl: The Southern Plains in the 1930s.* New York: Oxford: Oxford University Press, 1979.

_____ . *Dust Bowl: The Southern Plains in the 1930s*. New York: Oxford University Press, 1979.

_____ . *Nature's Economy: A History of Ecological Ideas*. Cambridge: Cambridge University Press, 1977.

Yarranton, G. A. "Plant Ecology: A Unifying Model." *Journal of Ecology* 57 (1969): 245-50.

Yoccoz, N. G. "Use, Overuse, and Misuse of Significance Tests in Evolutionary Biology and Ecology." *Bulletin of the Ecological Society of America* 72 (1991): 106-111.