Journal of Economic Entomology, FORUM

Ecoinformatics for integrated pest management: expanding the applied insect ecologist's tool-kit

Jay A. Rosenheim[1,6], Soroush Parsa[2], Andrew A. Forbes[1,3], William A. Krimmel[1], Yao Hua Law[1], Michal Segoli[1], Moran Segoli[1], Frances S. Sivakoff[1], Tania Zaviezo[1,4], and Kevin Gross[5]

[1]Department of Entomology, University of California, Davis, CA 95616.

[2]CIAT (Centro Internacional de Agricultura Tropical) A.A. 6713, Cali, Colombia.

[3]Current address: Department of Biology, University of Iowa, Iowa City, IA 52242.

[4]Departamento de Fruticultura y Enología, Pontificia Universidad Catolica Chile, Santiago 30622, Chile.

[5]Department of Statistics, North Carolina State University, Raleigh NC 27695.

[6]Corresponding author, e-mail: jarosenheim@ucdavis.edu.

ABSTRACT Experimentation has been the cornerstone of much of IPM research. Here we aim to open a discussion on the possible merits of expanding the use of observational studies, and in particular the use of data from farmers or private pest management consultants in 'ecoinformatics' studies, as tools that might complement traditional, experimental research. The manifold advantages of experimentation are widely appreciated: experiments provide definitive inferences regarding causal relationships between key variables, can produce uniform and high quality data sets, and are highly flexible in the treatments that can be evaluated. Perhaps less widely considered, however, are the possible disadvantages of experimental research. Using the yield-impact study to focus the discussion, we address some reasons why observational or ecoinformatics approaches might be attractive as complements to experimentation. A survey of the literature suggests that many contemporary yield-impact studies lack sufficient statistical power to resolve the small, but economically important, effects on crop yield that shape pest management decision-making by farmers. Ecoinformatics-based data sets can be substantially larger than experimental data sets, and therefore hold out the promise of enhanced power. Ecoinformatics approaches also address problems at the spatial and temporal scales at which farming is conducted, can achieve higher levels of 'external validity', and can allow researchers to efficiently screen a large number of variables during the initial, exploratory phases of research projects. Experimental, observational, and ecoinformatics-based approaches may, if used together, provide more efficient solutions to problems in pest management than can any single approach, used in isolation.

KEY WORDS Ecoinformatics, observational studies, statistical power, economic injury level, causal inference

INTEGRATED PEST management (IPM) research is highly diverse in the questions addressed and the research approaches employed. Some subdisciplines of IPM research rely heavily on observational studies, including for example research in the landscape ecology of insect herbivores and their natural enemies (e.g., Thies & Tscharntke 1999; Gardiner et al. 2009; Bahlai et al. 2010). Nevertheless, experimentation remains the foundation of most pest management research. The goal of this paper is to open a discussion on the possible utility of expanding the toolkit of the applied insect ecologist to include a greater role for observational studies, and in particular to evaluate critically the potential for ecoinformatics to contribute to our science.

What is ecoinformatics? Perhaps because the field is so new, usage of the term 'ecoinformatics' is not uniform (e.g., Recknagel 2006; Williams et al. 2006; Vos et al. 2006; Bekker et al. 2007; McIntosh et al. 2007; Sucaet et al. 2008; Hale and Hollister 2009), but ecoinformatics studies often: (1) use pre-existing data sets ('data mining') instead of data sets gathered by the researchers themselves; (2) integrate data sets from multiple sources to create a composite data set; (3) use observational data, rather than experimental data; (4) address ecological questions at a larger spatial and temporal scale than is typically feasible within an experimental framework; (5) use larger amounts of data than are typically feasible within an experimental framework; and (6) necessitate novel applications of data management, database design, and statistical analysis tools because of the large, observational, and often heterogeneous data sets involved. Thus, ecoinformatics is an interdisciplinary field in which computer scientists, statisticians, and ecologists work hand in hand to grapple with large-scale ecological questions.

Is there a relevant body of pre-existing data that can be mined by IPM researchers? We suggest that a bountiful opportunity to employ ecoinformatics exists in IPM, because private pest management consultants and farm staff generate large quantities of data on insect densities and crop performance as part of their routine, but extensive, sampling efforts in commercial agriculture. Insect scouting data can be combined with additional data streams from farmers, other private consultants (e.g., agronomy consultants), and governmental sources, including data on plant growth and performance, pesticide use, agronomic practices, and landscape context, to address a wide range of questions relevant to agricultural insect ecology.

We begin with a small survey of recently published studies to characterize the current state of research practices. We then review and discuss the most salient strengths of experimental research, followed by a consideration of some particular strengths of observational or ecoinformatics-based research that may allow them to complement traditional experimental work. Finally we provide a brief introduction to statistical tools that may be particularly useful for the analysis of observational studies. Our views have been influenced by our recent efforts to conduct observational studies (Rosenheim et al. 2006; Parsa 2010; Parsa et al. 2010) and to use ecoinformatics to address pest management problems in California cotton (Forbes and Rosenheim 2010; unpublished). We will allude to these experiences below.

Literature Survey: Studies of Pest Impact on Crop Yield. To make our discussion more focused and tangible, we propose to view the field of IPM through the lens of one particular type of study: the yield-impact study, in which the relationship between insect densities and crop yield is characterized. The yield impact study is one of the foundations of modern pest management programs, because it is used to estimate the economic injury level: the number of insects that reduces yield sufficiently that management intervention is economically advantageous (Pedigo 2002). We acknowledge, however, that other types of agricultural pest management research may employ quite different research methodologies. Our goal, then, is to ask whether observational and ecoinformatics-based approaches can contribute to progress in areas of IPM research that have traditionally relied heavily upon experimentation.

To describe current research practices within the community of IPM researchers, we reviewed all yield-impact studies conducted in the field or in greenhouses and published in *Journal of Economic Entomology* or *Environmental Entomology* between January 2007 and June 2010. Thirty-six papers satisfied our criteria for inclusion in the review, namely (i) that the study include a measure of crop yield in response to variation in densities of an herbivorous arthropod, and (ii) that the variation in herbivore densities either be natural or the result of an experimental manipulation of some kind, but not solely a response to different crop plant genotypes. We characterized each study using four basic descriptors: (1) was the study observational or experimental (we define a study as experimental if the researcher manipulated arthropod densities by applying a treatment to each experimental unit either randomly, or at least without regard to other traits expressed by that experimental unit); (2) were the data collected by the researcher or by other persons (e.g., farmers, consultants, etc.); (3) was the research conducted on a commercial farm or on a research farm; and (4) what was the size of each replicate plot (in m^2) within the overall study layout. In addition, to quantify the statistical power of those studies that employed an experimental approach (see below for details), we attempted to gather five further metrics for each study: (1) the mean and SD of crop yield observed in the treatment with the lowest level of herbivory (henceforth, the "arthropod-free control"); (2) the number of replicates for the arthropod-free control treatment; (3) the number of replicates for the treatment with the next lowest level of herbivory (henceforth, the "lowest damage treatment"); (4) the value of the crop (dollars/acre); and (5) the cost of a single application of the pesticide most commonly used to suppress the arthropod that was the focus of the paper (cost of the material plus the cost of the application; dollars/acre). Papers that did not report the needed crop yield data (mean and SD) or replicate numbers were excluded from further analysis. In cases where the authors did not provide estimates of crop value, we obtained these data from other sources, including primarily the USDA National Agricultural Statistics Service (http://www.nass.usda.gov/). Data on the current commonest pesticide use practices and costs were obtained either from each paper or, if not reported there, from university extension web-sites or from personal communications with specialists; the full data set with references is available from JAR.

The survey shows that experimentation is the dominant means by which researchers study the effects of herbivory on crop yield. Thirty-five of the 36 reviewed studies (97%) were experimental, with just the one remaining study (3%) employing an observational, correlative approach. Of the 27 studies that provided all the data needed to conduct the power analysis, data were collected by the researchers in all cases (100%); none of the studies involved mining data collected by non-researchers. Studies were usually conducted on experimental farms (22/27 studies, 81%) and much less frequently in the fields of cooperating farmers (3/27, 11%; in the remaining three studies, the location of the field plots was not specified). We discuss further these and other results of the literature survey below.

Strengths of experimental approaches/weaknesses of observational or ecoinformatics approaches

In this section, we summarize briefly views that we expect are already widely understood and assimilated within the research community regarding the manifold strengths of experimental science and the corresponding weaknesses of observational studies. We use the yield impact study as an exemplar to focus the discussion.

Experiments produce definitive inferences of causal relationships; observational studies cannot. Assume that in a well-replicated experiment, a researcher generates one or more treatments by manipulating some variable, A, while holding other conditions as nearly constant as possible; assigns those treatments randomly to experimental units; and then measures a response variable, B. If the response variable B differs significantly across treatments, then the experimenter can infer with a high degree of confidence that a change in A causes a change in B. This ability of experiments to reveal the causal structure of the environment is their most singular strength (e.g., Diamond 1983; Paine 2010). In contrast, when a researcher observes a correlation between natural, pre-existing variation in variables C and D, it is difficult to know whether the correlation reflects a causal influence of C on D, of D on C, or whether C and D are not causally related to each other at all, and instead are both influenced by some other variable(s) E, F, etc., which may or may not have been measured by the experimenter.

A yield impact experiment that employed only observational data but that attempted to infer a causal relationship between herbivore densities and crop yield could probably run afoul in several different ways, but two seem particularly likely. First, some herbivores may prefer to attack weak or stressed host plants (e.g., bark beetles; the ‘plant stress hypothesis’; White 1984; Mattson and Haack 1987; Huberty and Denno 2004), which are likely to produce less yield than vigorous, unstressed host plants, irrespective of herbivore load. Herbivores that prefer to attack low-vigor host plants are thus likely to be negatively correlated with crop yield, even if the damage that they generate actually has no effect on yield. In this case, it is instead the variable(s) that caused the plant stress in the first place that is the causal factor (e.g., for the bark beetle *Scolytus rugulosus* attacking almond trees in California, the causal agent for both decreased almond yield and increased bark beetle populations might be a soil-borne pathogen in the genus *Phytophthora*; University of California 2002). Second, other herbivores may prefer to attack particularly vigorously growing host plants (e.g., gall-inducing herbivores, cicadas; the ‘plant vigor hypothesis’; Price 1991; Cornelissen et al. 2008; Yang and Karban 2009). If vigorous plants are high yielding plants, then the result could be a spurious positive correlation between herbivore densities and crop yield or the masking or distortion of what could be a true underlying negative effect of herbivores on yield. Thus, purely observational data sets linking herbivore densities to crop performance must be approached with great caution, especially when the herbivore does not select host plants at random with respect to the host plant's yield potential, or else we must adopt some means of controlling for underlying variation in plant vigor.

By reducing between-replicate variation, experiments augment statistical power. Many experiments are conducted in a 'common garden' setting, in which all environmental conditions that might influence the dependent variable B are held as nearly constant as possible, except for the one variable, A, that is to be manipulated experimentally. By doing this, experimenters reduce the magnitude of unexplained variation, and thereby enhance the experiment's ability to resolve the influence of variable A on variable B. Moreover, many alternative experimental designs are available to reduce unexplained variation when common-garden experiments are unfeasible. For example, blocking is a familiar technique in which experimental units are grouped by a known source of variation that could impact the response, such as soil fertility. By manipulating the experimental variable within blocks, variability attributable to the external source cancels out, allowing a more direct assessment of the effect of the manipulated variable. Measurable differences in experimental units or environmental conditions can also be controlled statistically with regression designs. Regression designs require stronger assumptions than blocking designs (namely, that the effect of the measurable extraneous variation can be modeled mathematically), but return enhanced statistical power for detecting the effect of the manipulated variable when those assumptions are viable.

Experiments are flexible; in principle, any treatment can be generated; observational studies are limited to extant variation. Experiments are the ultimate intellectual playground in which researchers can attempt to implement any manipulation that they can imagine. In contrast, observational studies are restricted to conditions that actually occur in the field. This is not a profound observation, but it is one with important implications for using observational data to assess the relationship between herbivore densities and crop yield. In particular, if farmers manage a particular pest in a uniformly aggressive manner, maintaining its densities at low levels, then an observational study will be unable to explore the effects of higher densities of the herbivore on plant performance. Furthermore, the costs and benefits of any universally-adopted farming practice will be recalcitrant to study with purely observational approaches. For example, sulfur is applied to nearly 100% of all commercial grape production in California to suppress the fungal pathogen *Erisiphe necator* (powdery mildew); therefore, a strictly observational approach cannot be used to evaluate the hypothesis that sulfur exacerbates problems with *Tetranychus* spp. spider mites or *Erythroneura* spp. leafhoppers (Costello 2007; Jepsen et al. 2007), because there are virtually no sulfur-free vineyards with which sulfur-treated vineyards can be compared.

Data uniformity, completeness, and quality may be higher for data collected by researchers than for data used in ecoinformatics studies. Researchers who gather their own data have a high degree of control over the quality of their observations. Uniform data collection protocols, the option to measure all variables thought to be relevant to the question being addressed, and the ability to adjust sampling intensity to achieve the desired level of sampling precision are all available to the researcher. In contrast, data mining always involves giving up some of this control over data uniformity, completeness, and, possibly, quality. In the context of IPM research, private pest control consultants may use a variety of different sampling methods to estimate the density of a given focal pest species, creating challenges in integrating multiples sources of data into one composite data set. Research comparing different sampling methods may allow different types of data to be inter-converted (e.g., Musser et al. 2007), but such studies are not always available. In many cases, density estimates may be qualitative (e.g., densities may be recorded as "trace", "low", "moderate", or "high") rather than quantitative. The sampling effort efficiencies demanded by the highly competitive workplace may not always be compatible with research objectives, and variables not thought to be critical to immediate management decisions are often not measured, even if they may be needed in a research context. On the other hand, it should not be forgotten that consultants are professional arthropod samplers: their livelihoods depend on producing useful estimates of pest densities, and they often have more experience in sampling than even the most seasoned researcher.

Pest control consultants and farmers may not wish to share data. An absolute prerequisite of using ecoinformatics to address IPM research objectives is to establish a collaborative relationship with the consultants and farmers whose data will form the core of the ecoinformatics data set. There are two primary obstacles to establishing this collaboration. First, essentially all of the data typically needed for IPM research (e.g., insect densities, crop yield, pesticide use) are 'sensitive' for the persons who might provide those data. Consultants may be reluctant to divulge information about fields in which pest populations escaped control and generated substantial damage. Farmers are notoriously, and understandably, secretive about the yields that they obtain; yield data and details of agronomic practices may represent important competitive edges in the marketplace. Finally, pesticide use data are often very sensitive, due to the sometimes considerable blurring of the boundaries between legal use, consistent with labeled restrictions, and illegal use. Promises to treat all data confidentially may ameliorate these concerns, but rarely eliminate them. Second, requests for data sharing invariably impose a time burden on collaborating consultants and farmers; records must be located and organized, and sampling methods and recording practices must be explained in detail to the researcher. We have discovered that by working during the winter, when farmers and consultants are generally less pressured by immediate crop management responsibilities, it is often easier to secure active cooperation. Furthermore, we have found that the single most important element in securing active collaboration from farmers and consultants is to ensure that the ecoinformatics study addresses questions that they view as important to their livelihoods. In that way, farmers and consultants can expect a fair return on their very real investment in the conduct of the research. Finally, it is important to note that any time some farmers choose to participate in data sharing while others do not, it creates a possible filtering of the data set that may introduce various biases.

Weaknesses of experimental approaches/strengths of observational or ecoinformatics approaches

Given the many strengths of experimental science, as summarized above, it might seem strange indeed to consider alternative approaches to IPM research. In this section, however, we present views that may not be as widely considered within the research community regarding the limitations of experimental science and the corresponding strengths of observational or ecoinformatics-based studies. We again use the yield impact study as an exemplar to focus the discussion.

Traditional experimental designs may not have sufficient power; large ecoinformatics data sets may provide greater power. As noted above, experimenters may augment their ability to detect the effects of causal variable A on response variable B by holding as nearly constant as possible all other environmental variables (the 'common garden' approach). Nevertheless, we suggest that traditional agricultural experimentation may often fail to produce sufficiently precise estimates of key crop performance variables to guide many pest management decisions that farmers must make in their daily operations. The problem is that effects on yield that are small (perhaps too small to be resolved by traditional experimentation) may still be economically important to a farmer whose profit margin may be quite thin (e.g., see http://coststudies.ucdavis.edu/current.php). For example, a farmer who works with a 10% profit margin will be strongly motivated to avoid even a 2% loss of yield from herbivory, especially when that farmer can do so by applying an inexpensive pesticide. But, can we measure such small effects on yield?

We first explore the hypothesis that traditional experimentation may lack sufficient power by examining a case-study co-authored by one of us; in so doing, we lay out the methodology that we use below in a broader, literature-survey based test of the hypothesis. Rosenheim et al. (1997) examined the yield impact of the cotton aphid, *Aphis gossypii*, feeding on seedling upland cotton plants, *Gossypium hirsutum*, in

California. Cotton grown in California is not an unusually high value crop: mean yields in 2009 were 1,613 pounds/acre, and the average price received by growers was \$0.715 per pound, generating a crop value of \$1,153 per acre. A farmer faced with a potentially damaging aphid population on seedling cotton is faced with a simple decision: should I apply an insecticide or not? A single application of an insecticide commonly used to suppress aphids currently costs approximately \$18.25 per acre (\$8.50 per acre for the insecticide itself and \$9.75 per acre for the aerial application). To maximize profits, farmers should apply an insecticide only if the application cost is less than the value of the crop yield that would be sacrificed if the insect populations were not suppressed. Thus, assuming that a single application of insecticide will completely eliminate any potential effect of aphids on seedling cotton, farmers will maximize their profits by applying an insecticide if the aphids would otherwise cause a loss of (\$18.25/\$1153)% of yield, or 1.58%. Did the experiments reported by Rosenheim et al. (1997) have sufficient power to resolve effects of this size?

We can think of an idealized yield-impact study as including a key contrast between two treatments: an “arthropod-free” control, replicated n_1 times, and a “threshold damage” treatment, replicated n_2 times, that generates the amount of yield loss that corresponds to the point at which the profit-maximizing farmer would switch from not intervening to intervening to suppress pest densities (the “economic injury level”). Note that it is not a trivial challenge for the researcher to create this threshold damage treatment; before conducting preliminary trials, the insect density that produces this level of damage will generally be unknown. Furthermore, the function relating the intensity of herbivore damage to plant performance (the compensation function) is highly variable in form, and is frequently non-linear (Dyer et al. 1993; Huhta et al. 2003; Gao et al. 2008). As a result, whereas treatments that generate greater amounts of yield loss can help to define the complete form of the compensation function and can allow researchers to resolve statistically significant yield effects, they are largely uninformative regarding the yield effects of lower levels of damage. For the farmer, then, the key problem is to identify the economic injury level: at what pest density does the amount of protectable yield loss equal the cost of the pesticide application? To answer this question, we need to be able to resolve a statistically significant yield loss for the threshold damage treatment. In the simplest possible case, this yield loss can be evaluated as a *t*-test,

$$t_s = \frac{\overline{Y_1} - \overline{Y_2}}{\sqrt{\left(\frac{s_1^2}{n_2} + \frac{s_2^2}{n_1}\right)}}, \qquad (1)$$

where t_s is the critical *t*-value for a contrast with $n_1 + n_2 - 2$ degrees of freedom, $\overline{Y_1}$ is the mean yield in the arthropod-free control, $\overline{Y_2}$ is the mean yield in the threshold damage treatment, and s_1 and s_2 are the sample standard deviations observed for the two treatments. To be as generous as possible in evaluating the power of yield impact experiments, we can consider the test to be 1-tailed (i.e., excluding the possibility of overcompensation). Because not all studies include a “threshold damage” treatment (i.e., one corresponding closely to an amount of damage that represents the point at which a farmer’s optimal behavior switches from ‘don’t intervene’ to ‘intervene’), we can conservatively estimate $\overline{Y_1}$ and s_1 from the reported arthropod-free control treatment data and assume that $s_2 = s_1$. Equation (1) can then be rearranged to calculate

$$\frac{\bar{Y}_1 - \bar{Y}_2}{\bar{Y}_1} = \frac{t_s}{\bar{Y}_1} \cdot \sqrt{\left(\frac{s_1^2}{n_2} + \frac{s_2^2}{n_1}\right)}, \tag{2}$$

which is the proportional loss of yield that an experimenter could expect to detect 50% of the time, given the power of the study. Still smaller yield losses would be detected less than half the time. Note again that this calculation is generous in ascribing statistical power to the experiment, because a 50% probability of detecting an effect is already somewhat marginal.

Rosenheim et al. (1997) reported two experiments. In the first, the observed yield in the arthropod-free control, with n_1 = 10 replicates, was 2596 ± 91 grams (mean ± SE). Thus, $\bar{Y}_1$ = 2596 and s_1 = SE$\cdot \sqrt{n}$ = 288. The treatment with the lowest level of arthropod damage was also replicated 10 times (n_2 = 10), so we can imagine that a hypothetical treatment poised at the level of crop damage where the optimal decision would shift from not intervening to intervening would also have been replicated 10 times; thus t_s = 1.734, and we can let $s_2 = s_1$ = 288. With equation (2) we can then calculate the smallest proportional loss of yield that would be detected with 50% probability as

$$\frac{\bar{Y}_1 - \bar{Y}_2}{\bar{Y}_1} = \frac{223}{2596} = 0.0859, \tag{3}$$

or an 8.59% yield loss. The second experiment, for which $\bar{Y}_1$=2371, $s_1 = s_2$ = 282, n_1 = 20, n_2 = 10, and t_s = 1.701, generates an analogous estimate of a 7.84% yield loss. On average, then, the smallest yield loss that these experiments can reasonably expect to resolve is 8.22%. Alarmingly, this is approximately 5 times (8.22%/1.58% = 5.2) greater than the proportional yield loss at which a farmer should start applying an insecticide to suppress a damaging pest. To encapsulate this problem, we define a study's 'power ratio' as

$$\text{Power ratio} = \frac{\left(\begin{array}{l}\text{the smallest proportional yield loss that}\\ \text{can be detected with 50\% probability}\end{array}\right)}{\left(\begin{array}{l}\text{the smallest proportional yield loss that would}\\ \text{motivate a farmer to suppress the focal pest population}\end{array}\right)}. \quad (4)$$

Clearly, and in contrast to this first example, it will be highly desirable to conduct experiments that achieve power ratios < 1. Power ratios >1 suggest that a fundamental disconnect exists between the effect sizes that researchers can detect and the effect sizes that drive the pest management decisions of profit-maximizing farmers.

We are not the first to identify this possible problem with statistical power. Ragsdale et al. (2007) noted emphatically that, even working with a relatively low-value crop (soybeans), where the power problem should be less acute, the economic injury level was associated with a yield loss that was so small that it was "immeasurable". How widespread is this problem of insufficient power?

We used our survey of recently-published yield impact studies to try to address this question. Twenty-seven of the 36 studies surveyed presented the needed data on crop yield (mean plus some measure of variability). For crops with multiple harvests per year, crop value for just a single harvest was used. Any time the authors of the original studies collapsed observations across multiple experiments or treatments to produce larger sample sizes, we used these aggregate yield estimates to achieve the greatest possible statistical power. Many studies reported multiple experiments individually and did not collapse results; in these cases, we calculated a power ratio for each experiment, and then averaged across the different power ratio estimates to obtain a single observation per study.

Our survey suggests that the problem of insufficient power is a general one (Fig. 1); indeed, none of the 27 studies achieved a mean power ratio <1 (the lowest value was 1.49; see Table 1). If we look instead at the distribution of power ratios for each experiment reported within the 27 published studies, our sample size increases (N = 159), but the result is not much more encouraging: the median power ratio is 8.0 (range: 0.60-578.2), and only 4 of the 159 experiments (2.5%) achieved a power ratio < 1.0.

It appears then that experimental yield impact studies only very rarely have the statistical power needed to resolve the economic injury level, and thus to guide one of the most basic decisions that farmers must make in their daily pest management practices. How can this problem be overcome? We suggest four possible approaches. First, for at least a subset of the pests that directly attack the marketed portion of the crop ('direct pests'), it is possible to evaluate yield loss directly, by quantifying the damaged or destroyed portion of the crop. This may greatly ameliorate the power problem. For example, increased herbivory by the navel orangeworm, *Amyelois transitella*, on almond nuts may generate a very small loss of yield (say, 1%), representing a small 'signal' that may be lost in the abundant 'noise' generated by the many other factors that cause variation in almond yield (e.g., variation in soil quality, water or nutrient availability, pollinator efficacy, presence of pathogens, etc.). In contrast, even a similarly small absolute increase in the proportion of the harvested almond nuts bearing distinctive *A. transitella* feeding damage (e.g., increasing from 1% to 2%), as detected in the packing house, may be easier to resolve statistically, because whereas the ‘signal’ is still small, the ‘noise’ is reduced, since *A. transitella* is the sole source of such damage.

A second approach is to retain the same commitment to experimentation, but to increase the number of replicate plots. This suggestion is tempered by the recognition that feasibility concerns regularly constrain the number of replicates possible in any single experiment. However, an approach that is being used increasingly frequently and, we think, with excellent results, is to pool research effort across multiple workers, creating consortia of researchers capable of producing experiments that are heavily replicated across space and time (e.g., Ragsdale et al. 2007; Chapman et al. 2009; Johnson et al. 2009; Musser et al. 2009a,b). Because statistical power increases only as the square root of replicate number, however, in most cases very large increases in research effort are required to push the power ratio into the desired range (e.g., a 25-fold increase in replicate number is needed to bring the power ratio from 5.0 → 1.0). The huge labor and capital requirements of such extensive experimentation is the most significant obstacle to further adoption of this approach to augmenting power. Analyses that combine observations across different places and times may also sacrifice some of the advantages of experiments over observational studies discussed above. For example, some authors have created composite data sets by combined data across experiments and then employing regression analyses relating pest density to yield; such analyses do enhance power very substantially, but also sacrifice some of the interpretational rigor associated with experimental data.

A third possible approach again derives extra statistical power by pooling data across multiple experiments, but now in a strictly *post hoc* manner through formal meta-analysis. This differs from the creation of consortia of researchers in that the experiments to be pooled will generally have been performed by different researchers without any original coordination of effort. Meta-analysis is now used widely in biology, in large part because it effectively increases sample sizes by synthesizing data across multiple studies (Harrison in press), thereby decreasing the likelihood of failing to reject a null hypothesis (e.g., that a pest has no effect on yield), even when it is false (i.e., Type II error). In agricultural pest management, meta-analysis will be feasible only for pest-crop combinations that have been studied repeatedly.

A fourth possible means of realizing the needed statistical power is to seek out much larger data sets, capitalizing on the substantial data collection efforts made by the community of private consultants and farm employees who routinely scout fields, i.e., ecoinformatics. Ecoinformatics approaches, although still in their infancy, hold the promise of data sets that are orders of magnitude larger than those generated in a traditional experiment. Although assembling farmer- and consultant-derived data into a usable database can require a significant investment of time and labor, it can still be much more efficient than generating the data *de novo*.

The spatial and temporal scales of many experimental studies do not match the scales of commercial agriculture; ecoinformatics studies generally achieve the appropriate match. Experimental studies are generally performed in small research plantings, employing relatively small treatment plots. Our survey of published yield impact studies revealed a median plot size of just 36.9 m^2 (Table 1), roughly equivalent to a square plot 6 m on a side. Ecologists have long discussed the problems of extending experimental results observed at one spatial scale to another (Diamond 1983; Addicott et al. 1987; Willis and Whittaker 2002; Paine 2010). In the case of IPM research, this problem is likely to be acute, because the difference in spatial scale may be large (often ≈2 orders of magnitude). We offer one example of the problems that may be encountered in attempting to scale up. One of the commonest problems encountered in agricultural pest management is the potential of broad-spectrum insecticide applications to elicit pest resurgences or secondary pest outbreaks as a result of suppressing natural enemy populations (Hardin et al. 1995). Experimentation examining pest suppression with pesticides in small research plots may be unlikely to reveal the full scope of possible problems with resurgences or secondary pest outbreaks, because it is easy for natural enemies to move just the handful of meters required to re-colonize sprayed plots from adjacent unsprayed plots. In contrast, when natural enemy populations in a large commercial field are suppressed by a pesticide, re-colonization requires beneficial insects to travel much farther, and thus may take too long to prevent pest population eruptions.

The problem with temporal scale is different. Most yield impact studies conducted with annual crops are indeed performed at the appropriate temporal scale (a whole cropping cycle). But yield impact studies for perennial crops may require experimental manipulations to be maintained for several years to quantify the cumulative effects of herbivore stress, and then crop performance must be observed for years following the removal of herbivory to assess the possibility for lagged effects. Such multi-year yield-impact studies have been successfully conducted (Welter et al. 1989, 1991; Hare et al. 1999; Fournier et al. 2006), but the requirement for multiple years of experimentation makes the work very costly. These costs discourage researchers from updating economic injury levels as agronomic practices change (e.g., introductions of new crop cultivars) and limit yield impact studies to only a small handful of the most important pests. It is probably not just a coincidence that none of the 36 studies reviewed in our survey dealt with a perennial crop.

Observational studies performed in farmers' fields and ecoinformatics-based approaches largely avoid these problems of spatial and temporal scale. When data are collected in the real commercial farming setting, there is no need to translate to a different spatial scale. Ecoinformatics approaches also hold out the hope of capturing quickly and efficiently multiple years of data on pest densities and performance of both annual and perennial crops when cooperating consultants and farmers have adequate record keeping. Although record-keeping practices vary, our experience has been that many consultants do retain their pest monitoring data for several years. The ecoinformatics approach will not be a panacea for all problems of temporal scale; for example, a crop rotation scheme that led to gradual soil acidification and the establishment of an acid-loving soil-borne pathogen did not emerge until years 40-80 of a long-term experiment conducted by scientists at the Rothamsted Agricultural Research Station (Denison 2011). Such problems are, hopefully, exceptional in the context of arthropod management.

The narrowly controlled environmental conditions of experimental studies give strong 'internal validity,' but may restrict the ability to extend conclusions to situations of different environmental conditions (i.e., limited 'external validity'). As noted above, researchers often augment the statistical power of their experiments by holding environmental conditions as nearly constant as possible. Although this approach has obvious merits, it does raise the question of whether or not the conclusions derived from the experiment are relevant to farming operations that are conducted under other conditions (e.g., different crop cultivars, soil types, microclimates, or agronomic practices; presence of other members of a frequently speciose food web centered on the crop plant, including other herbivores, plant pathogens, omnivores, and predators). The spatial and temporal scale issues discussed above are just one expression of this more general problem. The importance of choosing research methods that recognize the trade-off between internal and external validity has been discussed in diverse fields (e.g., community ecology: Diamond 1983; Miller 1986; economics: Roe and Just 2009).

Of course, repeating experiments at different locations and at different times helps to build confidence that conclusions are more broadly relevant. But simply repeating experiments does not solve all aspects of this problem. For example, 22 out of the 25 (88%) of our surveyed yield-impact studies that specified where the experiments were conducted were performed in research farms, with only the remaining 3 studies (12%) performed in cooperating farmers' fields (Table 1). This may reflect the prevalence within the journals we surveyed of studies performed in North America, where research farms are commonplace; in other regions of the world, research in commercial farmers' fields may be more common. Although research farms do offer potential advantages for experimentation, research farms also differ in many ways from the commercial setting. Farmers are often reluctant to adopt pest management recommendations derived from small experiments performed on research farms; this is a major reason why cooperative extension specialists often establish demonstration plots in farmers' fields – to show farmers that practices actually work when applied in the commercial setting.

As noted by Jiménez et al. (2009), observational studies conducted in commercial fields and ecoinformatics-based data sets can largely avoid these problems, because the data can be collected from many commercial fields. With careful planning, the data can reflect a representative range of the diverse conditions under which the crop is farmed. This purposeful 'heterogenization' of the data set (see Richter et al. 2009, 2010) can increase the confidence with which farmers view a study's conclusions.

Observational or ecoinformatics-based approaches may be particularly valuable as a means of screening a large number of potentially important variables during the early, exploratory phase of a research project. IPM research often involves highly focused research questions; the yield-impact study that has guided this opinion piece is one such example, in which the relationship between just two variables (herbivore density and crop yield) is to be examined. But, in some cases, IPM research may begin with more open-ended or ill-defined questions, which necessitate an initial, highly exploratory phase of research in which a large number of candidate variables are screened to identify a smaller set of variables that is amenable to experimental analysis (e.g., Jiménez et al. 2009). Whereas experimental designs capable of screening a larger number of variables do exist (e.g., fractional factorial designs), they necessitate a larger-than-usual number of experimental plots, may be taxing because the experimenter may need to devise novel means of manipulating many variables, and have limited abilities to explore interactions between multiple factors. Observational and ecoinformatics-based studies may be particularly valuable during the early stages of a highly exploratory research program, when the main goal is to shorten the list of variables and generate hypotheses for further, more narrowly focused testing. In this regard, ecoinformatics data sets that represent a large range of commercial farming conditions also provide enhanced opportunities to screen the effects of multiple variables on a pest-crop interaction.

One example should make clear the potential complementarity of an initial observational phase of research followed by a subsequent, more narrowly focused experimental phase of research. Cotton farmers in California have long noted that the short-term appearance of crop damage produced by *Lygus hesperus* feeding on cotton is highly enigmatic: in some fields with many *Lygus*, little damage (the shedding of young flower buds) is seen, whereas in other fields with few *Lygus*, high damage is observed. Why? The list of possible explanations was dauntingly large; under the broad headings of (i) observer error; (ii) variable insect behavior; (iii) variable plant response; and (iv) crop damage produced by some other insect; 23 variables were screened in an observational study conducted in farmers' fields (Rosenheim et al. 2006). The observational study allowed us to cast a wide net, and suggested a completely unexpected underlying mechanism for enigmatic crop damage, namely that it was the cotton plant's phosphorus content, itself a reflection of the field's crop rotation history, that controlled the plant's response to *Lygus* feeding damage (i.e., the key effect was an interaction of phosphorus and *Lygus* herbivory). Subsequent manipulative experimentation confirmed a direct causal role for phosphorus (Forbes and Rosenheim, unpubl. ms.). Because manipulating phosphorus proved to be very difficult (a large field experiment failed to establish the desired nutrient level treatments; it took three successive tries in the greenhouse to produce the right nutrient and damage treatments), this result likely would never have been obtained if all 23 variables had to be explored experimentally from the start. With no reason to suspect a role for phosphorus (no such suggestion existed in the extensive literature on flower bud abscission in cotton; Addicott 1982; Weir et al. 1996), such an ambitious set of experiments to screen for a phosphorus effect would have been unthinkable. Thus, while the observational study alone was not sufficient in this case to generate any confidence that the correlation was real or reflected a causal relationship, the combination of observational and experimental approaches answered a long-standing question that otherwise would likely have remained a mystery.

Researchers and farmers may use different sampling methods, and translating research results into decision tools that farmers can use may be challenging. In each of the 27 studies that provided the data needed for the power analysis, all data were collected by the researchers themselves. As discussed above, when researchers gather their own data, they may secure the benefits of high data uniformity and quality. However, it is also often the case that researchers use sampling methods that differ from those used in commercial pest scouting operations. In such cases, it may be difficult to ‘translate’ research-based recommendations, generated with one sampling methodology, to a farmer-ready decision tool that will be implemented with a different sampling method. This is not an insurmountable problem, but is one that may mandate additional research effort. Ecoinformatics-based approaches, on the other hand, use farmer-generated data to produce decision rules that are immediately ready to be implemented in the same ‘language’ as the original data set; nothing should be lost in translation.

Statistical tools for observational and ecoinformatics data sets

As we have seen, observational data can be used to elucidate and quantify relationships between key variables in IPM. However, merely detecting an association in observational data provides no evidence that the association is causal, that is, that variation in one variable generates variation in the other. This limitation of observational data is broadly appreciated. Despite this limitation, observational data can still provide a basis for scientific learning, especially when observational studies are coupled with experiments. The example of phosphorous content mediating *Lygus* damage to cotton described above illustrates this possibility. Thus, observational data complement experimental data, and together the two can foster learning about causal relationships in IPM.

However, and perhaps surprisingly, causal learning with observational data does not always have to be informal. In fact, there is a restricted set of circumstances under which observational data themselves can be used to draw inferences about cause-and-effect relationships in a mathematically rigorous way. These circumstances, and the statistical methods that can be used for causal inference when they prevail, are the topics to which we now turn. Statistical methods for drawing causal inferences from observational data have been developed largely in the context of disciplines that study human welfare, namely the behavioral sciences (particularly economics: Rosenbaum 2002; Imbens and Wooldridge 2009; Gangl 2010) and public health (Little and Rubin 2000; Jewell 2004). In these settings, the notion of experimentally manipulating the putative causal variable of interest (e.g., wages, or exposure to an environmental toxicant) is either unfeasible, unethical, or both. Consequently, investigators in these fields have pioneered the development of methodologies for eliciting causal inferences from observational data. We suggest that some of these methods can be fruitfully applied to observational data in the natural sciences as well.

A comprehensive survey of statistical methods for causal inference is beyond the scope of this article. Instead, our goal in this section is to discuss general insights that have emerged from this literature, and to provide references that may serve as a gateway for the interested reader. Among the references cited in this section, Imbens and Wooldridge (2009) provide a particularly readable and comprehensive review of the field. We plan to present a more detailed case study of causal inference in IPM in a future contribution.

The key insight to emerge from the causal-inference literature is that causal inference from observational data is only possible if covariates are available that eliminate confounding between the putative cause and response variables. This 'no unmeasured confounders' condition is perhaps not surprising, and it is also not necessarily discouraging – an understanding of the conditions required for formal causal inference does not prohibit informal learning under any circumstance, and indeed opens the door to formal causal inference in those scenarios where the condition is met. Evaluating the 'no unmeasured confounders' assumption also requires clearly articulating the conditions under which a covariate qualifies as a confounder. In short, a covariate is a confounder if it is causally associated with both the putative causal variable and the putative response (Jewell 2004). For example, in a yield-impact study, plant vigor is a confounder if vigor either attracts or deters arthropod herbivores and simultaneously impacts yield through other pathways unrelated to arthropod feeding. Jewell (2004) describes graphical approaches that can be used to identify confounding variables.

Clearly, evaluating the 'no unmeasured confounders' condition requires a deep and thorough knowledge of the system under study. Although this condition will surely need to be evaluated on a case-by-case basis, it is conceivable to us that some IPM questions may lend themselves to satisfying this condition more naturally than others. In particular, identifying confounders may be more feasible when the number of recognized management options available to IPM practitioners is small, and when managers or farmers record and make available the scouting information (e.g., arthropod densities, weather conditions) that they use to decide which of these options to pursue.

If the no unmeasured confounders condition is met, methods exist for drawing causal inferences about the relationship between the causal variable and the response. We provide the briefest of introductions to two of these methods here, and point the interested reader to references that provide a more thorough description. A versatile method for eliciting causal relationships is multiple regression. Here, one builds a regression model in which the putative cause, the confounder(s) and their statistical interactions are included as predictors in the regression model. Multiple regression models are attractive when the number of confounders is large, and/or when the confounders are continuous variables. A subtlety here is that the causal effect of the putative causal variable on the response is not in general equal to the partial regression coefficient associated with the causal effect. Instead, the causal effect is estimated by evaluating the fitted regression model for different values of the causal variable and all the observed values of the confounders. Regression methods can also be used when the causal effect depends on the value of one or more covariates. Regression methods for causal inference are described in Imbens and Wooldridge (2009).

A second but related approach entails the use of propensity scores (Rosenbaum and Rubin 1983). Propensity scores are especially useful when the putative causal variable is binary, such as whether or not a particular management intervention was used. Use of propensity scores entails two stages of modeling. In the first stage, one builds a statistical model in which the confounders serve as predictors and the putative causal variable serves as the response. Propensity scores are the fitted values from that model, and reflect the information about the treatment assignment contained in the confounders. A variety of estimators are then available to quantify the causal effect of the treatment, either by stratifying on or weighting by the propensity score. Recent reviews of propensity score methods can be found in D'Agostino (1998) and Lunceford and Davidian (2004).

Consideration of statistical methods for causal inference also brings to light useful principles that can inform the design of an observational study. First, the 'no unmeasured confounders' assumption clearly limits the type of questions for which observational data can be used to measure causality directly. In particular, 'no unmeasured confounders' demands that the investigator possess sufficient expertise to knowledgably assess whether or not the variables in hand capture all possible sources of confounding. Second, a 'greedy' approach in which one amasses as much data as possible and hopes that learning will ensue is not necessarily wise or efficient. Intelligent construction of observational data sets requires that the data gathered span the range of interesting variability for both the causal variable of interest and any confounders. For example, in yield-impact studies, selection of an appropriate 'control' that allows one to quantify yield when the arthropod is absent (or at least minimally present) is vital. Haphazard or random collection of observational data does not ensure that a suitable control will be included, and offers no benefit equivalent to random assignment of treatments in controlled experiments. Thus, much like experimental studies, observational studies also benefit from careful forethought in the planning stages, and well-constructed observational data sets will strengthen the analyst's ability to draw causal inferences about the IPM system under study.

Conclusions

The advantages of experimental research are well appreciated by applied insect ecologists; foremost among these is the ability to make definitive inferences regarding causal relationships between variables. Nevertheless, our analysis suggests that experimental science, like any approach to science, has both strengths and weaknesses. We have argued that a key weakness of experimentation in agricultural pest management research is the frequent lack of sufficient statistical power to resolve the small but economically important yield effects that dictate farmer pest management decisions. Observational approaches to science, while clearly at a disadvantage in determining causal relationships, have strengths that can largely complement the weaknesses of experimental science. In particular, ecoinformatics-based approaches can produce data sets that are substantially larger than typical experimental data sets, producing opportunities for improved power. Observational and ecoinformatics studies can also more readily address questions at the true spatial and temporal scale of commercial agriculture and can embrace a large range of the natural variation in commercial farming conditions. For these reasons, observational studies are growing in their importance within IPM research (e.g., Rochester et al. 2002; Carrière et al. 2004; Cattaneo et al. 2006; Gardiner et al. 2009; Jiménez et al. 2009; de Valpine et al. 2010). A vigorous analysis and discussion of the relative strengths and weaknesses of different research approaches can, we suggest, encourage researchers to combine the complementary strengths of different approaches (Diamond 1983), thereby helping to accelerate progress in IPM research and the agricultural sciences more broadly.

Acknowledgements

For helpful feedback on an earlier draft of the manuscript, we thank Irina Shapiro. Funding for this project was provided by the California State Support Board of Cotton Incorporated and by the USDA-NRICGP, grant 2006-01761.

References Cited

Addicott, F. T. 1982. Abscission. University of California Press, Berkeley, CA.

Addicott, J. F., J. M. Aho, M. F. Antolin, D. K. Padilla, J. S. Richardson, and D. A. Soluk. 1987. Ecological neighborhoods: scaling environmental patterns. Oikos 49:340-346.

Bahlai, C. A., S. Sikkema, R. H. Hallett, J. Newman, and A. W. Schaafsma. 2010. Modeling distribution and abundance of soybean aphid in soybean fields using measurements from the surrounding landscape. Environ. Entomol. 39:50-56.

Bekker, R. M., E. van der Maarel, H. Bruelheide, and K. Woods. 2007. Long-term datasets: from descriptive to predictive data using ecoinformatics. J. Veg. Sci. 18:458-462.

Carrière, Y., P. Dutilleul, C. Ellers-Kirk, B. Pedersen, S. Haller, L. Antilla, T. J. Dennehy, and B. E. Tabashnik. 2004. Sources, sinks, and the zone of influence of refuges for managing insect resistance to Bt crops. Ecol. Appl. 14:1615-1623.

Cattaneo, M. G., C. Yafuso, C. Schmidt, C.-Y. Huang, M. Rahman, C. Olson, C. Ellers-Kirk, B. J. Orr, S. E. Marsh, L. Antilla, P. Dutilleul, and Y. Carrière. 2006. Farm-scale evaluation of the impacts of transgenic cotton on biodiversity, pesticide use, and yield. Proc. Natl. Acad. Sci. USA 103:7571-7576.

Chapman, A. V., T. P. Kuhar, P. B. Schultz, T. W. Leslie, S. J. Fleischer, G. P. Dively, and J. Whalen. 2009. Integrating chemical and biological control of European corn borer in bell pepper. J. Econ. Entomol. 102:287-295.

Cornelissen, T., G. W. Fernandes, and J. Vasconcellos-Neto. 2008. Size does matter: variation in herbivory between and within plants and the plant vigor hypothesis. Oikos 117:1121-1130.

Costello, M. J. 2007. Impact of sulfur on density of *Tetranychus pacificus* (Acari: Tetranychidae) and *Galendromus occidentalis* (Acari: Phytoseiidae) in a central California vineyard. Exp. Appl. Acarol. 42:197-208.

D'Agostino, R.B. 1998. Propensity score methods for bias reduction in the comparison of a treatment group to a non-randomized control group. Stat. Med. 17:2265-2281.

de Valpine, P., K. Scranton, and C. P. Ohmart. 2010. Synchrony of population dynamics of two vineyard arthropods occurs at multiple spatial and temporal scales. Ecol. Appl. 20:1926-1935.

Denison, R. F. 2011. Darwinian agriculture: where does nature's wisdom lie? Princeton University Press, Princeton, NJ.

Diamond, J. M. 1983. Laboratory, field and natural experiments. Nature 304:586-587.

Dyer, M. I., C. L. Turner, and T. R. Seastedt. 1993. Herbivory and its consequences. Ecol. Appl. 3:10-16.

Forbes, A. A., and J. A. Rosenheim. Phosphorus nutrition modulates plant responses to insect herbivore damage. (unpubl. ms.)

Fournier, V., J. A. Rosenheim, J. Brodeur, J. M. Diez, and M. W. Johnson. 2006. Multiple plant exploiters on a shared host: testing for nonadditive effects on plant performance. Ecol. Appl. 16:2382-2398.

Gangl, M. 2010. Causal inference in sociological research. Ann. Rev. Soc. 36:21-47.

Gao, Y., D. Wang, L. Ba, Y. Bai, and B. Liu. 2008. Interactions between herbivory and resource availability on grazing tolerance of *Leymus chinensis*. Environ. Exp. Bot. 63:113-122.

Gardiner, M. M., D. A. Landis, C. Gratton, C. D. DiFonzo, M. O'Neal, J. M. Chacon, M. T. Wayo, N. P. Schmidt, E. E. Mueller, and G. E. Heimpel. 2009. Landscape diversity enhances biological control of an introduced crop pest in north-central USA. Ecol. Appl. 19:143-154.

Hale, S. S., and J. W. Hollister. 2009. Beyond data management: how ecoinformatics can benefit environmental monitoring programs. Environ. Monit. Assess. 150:227-235.

Hare, J. D., M. Rakha, and P. A. Phillips. 1999. Citrus bud mite (Acari: Eriophyidae): an economic pest of California lemons? J. Econ. Entomol. 92:663-675.

Hardin, M. R., B. Benrey, M. Coll, W. O. Lamp, G. K. Roderick, and P. Barbosa. 1995. Arthropod pest resurgence: an overview of potential mechanisms. Crop Prot. 14:3-18.

Huberty, A. F., and R. F. Denno. 2004. Plant water stress and its consequences for herbivorous insects: a new synthesis. Ecology 85:1383-1398.

Huhta, A.-P., K. Hellström, P. Rautio, and J. Tuomi. 2003. Grazing tolerance of *Gentianella amarelle* and other monocarpic herbs: why is tolerance highest at low damage levels? Plant Ecol. 166:49-61.

Imbens, G. W., and J. M. Wooldridge. 2009. Recent developments in the econometrics of program evaluation. J. Econ. Lit. 47:5-86.

Jepsen, S. J., J. A. Rosenheim, and M. E. Bench. 2007. The effect of sulfur on biological control of the grape leafhopper, *Erythroneura elegantula*, by the egg parasitoid *Anagrus erythroneurae*. BioControl 52:721-732.

Jewell, N.P. 2004. Statistics for epidemiology. Chapman & Hall, Boca Raton, FL, USA.

Jiménez, D., J. Cock, H. F. Satizábal, M. A. Barreto, A. Pérez-Uribe, A. Jarvis, and P. Van Damme. 2009. Analysis of Andean blackberry (*Rubus glaucus*) production models obtained by means of artificial neural networks exploiting information collected by small-scale growers in Colombia and publicly available meteorological data. Comp. Electron. Agric. 69:198-208.

Johnson, K. D., M. E. O'Neal, D. W. Ragsdale, C. D. DiFonzo, S. M. Swinton, P. M. Dixon, B. D. Potter, E. W. Hodgson, and A. C. Costamagna. 2009. Probability of cost-effective management of soybean aphid (Hemiptera: Aphididae) in North America. J. Econ. Entomol. 102:2101-2108.

Little, R.J., and D.B. Rubin. 2000. Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. Ann. Rev. Pub. Health 21:121-145.

Lunceford, J.K., and M. Davidian. 2004. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. Stat. Med. 23:2937-2960.

Mattson, W. J., and R. A. Haack. 1987. The role of drought in outbreaks of plant-eating insects. BioScience 37:110-118.

McIntosh, A. C. S., J. B. Cushing, N. M. Nadkarni, and L. Zeman. 2007. Database design for ecologists: composing core entities with observations. Ecol. Inform. 2:224-236.

Miller, J. 1986. Manipulations and interpretations in tests for competition in streams: "controlled" vs "natural" experiments. Oikos 47:120-123.

Musser, F., S. Stewart, R. Bagwell, G. Lorenz, A. Catchot, E. Burris, D. Cook, J. Robbins, J. Greene, G. Studebaker, and J. Gore. 2007. Comparison of direct and indirect sampling methods for tarnished plant bug (Hemiptera: Miridae) in flowering cotton. J. Econ. Entomol. 100:1916-1923.

Musser, F. R., A. L. Catchot, S. D. Stewart, R. D. Bagwell, G. M. Lorenz, K. V. Tindall, G. E. Studebaker, B. R. Leonard, D. S. Akin, D. R. Cook, and C. A. Daves. 2009a. Tarnished plant bug (Hemiptera: Miridae) thresholds and sampling comparisons for flowering cotton in the midsouthern United States. J. Econ. Entomol. 102:1827-1836.

Musser, F. R., G. M. Lorenz, S. D. Stewart, R. D. Bagwell, B. R. Leonard, A. L. Catchot, K. V. Tindall, G. E. Studebaker, D. S. Akin, D. R. Cook, and C. A. Daves. 2009b. Tarnished plant bug (Hemiptera: Miridae) thresholds for cotton before bloom in the midsouth of the United States. J. Econ. Entomol. 102:2109-2115.

Paine, R. T. 2010. Macroecology: does it ignore or can it encourage further ecological syntheses based on spatially local experimental manipulations? Amer. Nat. 176:385-393.

Parsa, S. 2010. Native herbivore becomes key pest after dismantlement of a traditional farming system. Am. Entomol. (in press).

Parsa, S., R. Ccanto, and J. A. Rosenheim. 2010. Resource concentration dilutes a key pest in indigenous potato agriculture**.** Ecol. Appl. (in press).

Pedigo, L. P. 2002. Entomology and pest management, 4th ed. Prentice Hall, New Jersey.

Price, P. W. 1991. The plant vigor hypothesis and herbivore attack. Oikos 62:244-251.

Ragsdale, D. W., B. P. McCornack, R. C. Venette, B. D. Potter, I. V. MacRae, E. W. Hodgson, M. E. O'Neal, K. D. Johnson, R. J. O'Neil, C. D. DiFonzo, T. E. Hunt, P. A. Glogoza, and E. M. Cullen. 2007. Economic threshold for soybean aphid (Hemiptera: Aphididae). J. Econ. Entomol. 100:1258-1267.

Recknagel, F. 2006. Editorial. Ecol. Inform. 1:1.

Richter, S. H., J. P. Garner, and H. Würbel. 2009. Environmental standardization: cure or cause of poor reproducibility in animal experiments? Nature Meth. 6:257-261.

Richter, S. H., J. P. Garner, C. Auer, J. Kunert, and H. Würbel. 2010. Systematic variation improves reproducibility of animal experiments. Nature Meth. 7:167-168.

Rochester, W. A., M. P. Zalucki, A. Ward, M. Miles, and D. A. H. Murray. 2002. Testing insect movement theory: empirical analysis of pest data routinely collected from agricultural crops. Comput. Electron. Agric. 35:139-149.

Roe, B., and D. Just. 2009. Internal and external validity in economics research: tradeoffs between experiments, field experiments, natural experiments, and field data. Am. J. Agric. Econ. 91:1266-1271.

Rosenbaum, P., and D. Rubin. 1983. The central role of the propensity score in observational studies for causal effects. Biometrika 70:41-55.

Rosenheim, J. A., K. Steinmann, G. A. Langellotto, and A. G. Zink. 2006. Estimating the impact of *Lygus hesperus* on cotton: the insect, plant, and human observer as sources of variability. Environ. Entomol. 35:1141-1153.

Rosenheim, J. A., L. R. Wilhoit, P. B. Goodell, E. E. Grafton-Cardwell, and T. F. Leigh. 1997. Plant compensation, natural biological control, and herbivory by *Aphis gossypii* on pre-reproductive cotton: the anatomy of a non-pest. Entomol. Exper. Appl. 85:45-63.

Sucaet, Y., J. Van Hemert, B. Tucker, and L. Bartholomay. 2008. A web-based relational database for monitoring and analyzing mosquito population dynamics. J. Med. Entomol. 45:775-784.

Thies, C., and T. Tscharntke. 1999. Landscape structure and biological control in agroecosystems. Science 285:893-895.

University of California. 2002. Integrated pest management for almonds, 2nd ed. University of California Division of Agriculture and Natural Resources, Publ. 3308, Oakland, California.

Vos, M., L. E. M. Vet, F. L. Wäckers, J. J. Middelburg, W. H. van der Putten, W. M. Mooij, C. H. R. Heip, and E. van Donk. 2006. Infochemicals sructure marine, terrestrial and freshwater food webs: implications for ecological informatics. Ecol. Inform. 1:23-32.

Weir, B. L., T. A. Kerby, K. D. Hake, B. A. Roberts, and L. J. Zelinski. 1996. Cotton fertility, pp. 210-227. *In* S. J. Hake, T. A. Kerby, and K. D. Hake [eds], Cotton production manual. University of California, Division of Agricultural and Natural Resources, Publ. 3352, Oakland, CA.

Welter, S. C., P S. McNally, and D. S. Farnham. 1989. Effect of Willamette mite (Acari: Tetranychidadae) on grape productivity and quality: a reappraisal. Environ. Entomol. 18:408-411.

Welter, S. C., R. Freeman, and D. S. Farnham. 1991. Recovery of 'Zinfandel' grapevines from feeding damage by Willamette spider mite (Acari: Tetranychidae): implications for economic injury level studies in perennial crops. Environ. Entomol. 20:104-109.

White, T. C. R. 1984. The abundance of invertebrate herbivores in relation to the availability of nitrogen in stressed food plants. Oecologia 63:90-105.

Williams, R. J., N. D. Martinez, and J. Golbeck. 2006. Ontologies for ecoinformatics. J. Web Seman. 4:237-242.

Willis, K. J., and R. J. Whittaker. 2002. Species diversity – scale matters. Science 295:1245-1248.

Yang, L. H., and R. Karban. 2009. Long-term habitat selection and chronic root herbivory: explaining the relationship between periodical cicada density and tree growth. Am. Nat. 173:105-112.

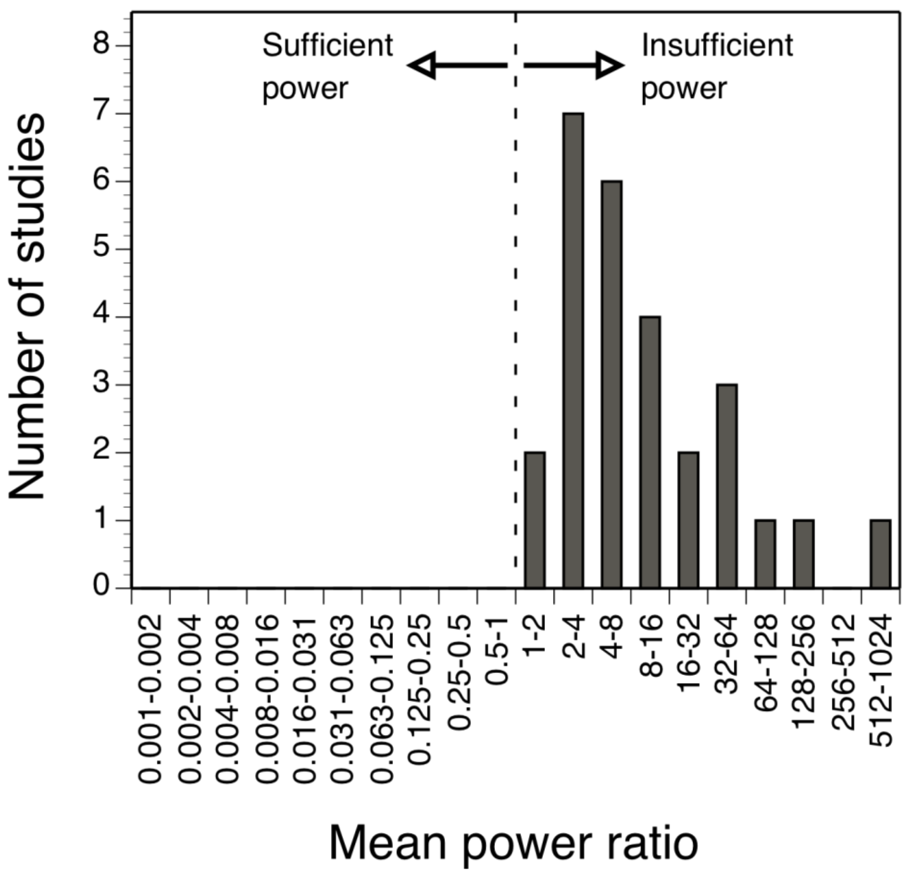
Table 1. Summary statistics for the survey of recently published studies (2007-2010) examining the relationship between pest density and crop yield (*N* = 36 total studies, 27 of which provided the data needed to calculate a power ratio). For each of the variables described in the table, each study provided a single observation (when multiple experiments were reported in a single publication, the mean value of the variable across the experiments is reported).

A. Categorical variables		%
Type of study:	35/36 experimental	97.2
	1/36 observational	2.8
Persons responsible for data collection:	27/27 researchers	100.0
	0/27 others	0.0
Location of field trial:	2/27 not stated	74.1
	3/27 commercial farms	11.1
	22/27 experimental farm	81.5

B. Continuous variables	Median (Mean; SD)	Range
Plot size (m^2)	36.9 (486.7; 1058.8)	2.0-4000
Number of replicates for lowest herbivory treatment	4 (11.7; 21.1)	2-90
Number of replicates for next lowest herbivory treatment	4 (11.8; 21.1)	2-90
Crop value (dollars/acre)	555 (4,348; 11,649)	145.7-56,089
Cost of a single pesticide application	14.5 (18.2; 18.5)	6.0-102.2

(dollars/acre)		
Smallest proportional yield loss detectable with 50% probability	0.109 (0.186; 0.173)	0.042-0.691
Smallest proportional yield loss that would motivate a farmer to suppress the pest population	0.020 (0.021; 0.014)	0.00037-0.048
Power ratio	4.69 (41.6; 113.8)	1.49-578.2

Figure 1. Survey of recently published studies (2007-2010; $N = 27$) examining the relationship between pest density and crop yield. For each study the mean power ratio was calculated as (the smallest proportional yield loss that would be detected with a probability of 50%)/(the smallest proportional yield loss that would motivate a farmer to suppress the focal pest population). Each study contributed one observation (power ratios were averaged across experiments for studies reporting multiple experiments). This ratio should be <1 for the study to have sufficient power to resolve the economic injury level for the pest, and thus to guide pest management practices; however, none of the studies achieved this desired level of statistical power.


Sufficient power
Insufficient power
Number of studies
0
1
2
3
4
5
6
7
8
0.001-0.002
0.002-0.004
0.004-0.008
0.008-0.016
0.016-0.031
0.031-0.063
0.063-0.125
0.125-0.25
0.25-0.5
0.5-1
1-2
2-4
4-8
8-16
16-32
32-64
64-128
128-256
256-512
512-1024
Mean power ratio