Butterfly Conservation Management in Midwestern Open Habitats Part 6: Surveying and monitoring butterflies

Chapter · January 2011

1 author:

Ann Swengel
no institution

Butterfly Conservation Management in Midwestern Open Habitats

Part 6: Surveying and monitoring butterflies
by Ann B. Swengel

I advise at least 50% of the resources for a habitat conservation project be devoted to surveying and monitoring the species and sites involved. Field data on what is in a site and where, and how those species are faring, are essential to understanding how habitat management is affecting them. Both butterfly data from other sites historically and currently supporting those species, and vegetative data from other sites of similar habitat representing a range of management approaches, can be invaluable for understanding the results at the sites you're focusing on.

Decisions about how to survey and monitor involve an examination of tradeoffs. The goal is to identify how to get the most of what you want within the constraints you have. Surveying is first about determining what occurs in your site (an "inventory" of species). If done over enough years, surveying can also monitor how those species are faring in your site over time. The tradeoffs concern such issues as who is available to do the field work, how experienced they are, what their motivation(s) are (fun, social event, scientific data, etc.), what you and they are willing to do, how many sites need to be surveyed, how many and which species need to be covered effectively, whether you already have a good idea what's there or are starting from scratch, how much budget is available, how much experience and expertise at data management and analysis are available, and what you want to learn. I can sure testify that there is a tradeoff between maximizing enjoyment at the moment (as a hobby) compared to collecting and preserving the data on the scale needed to get enough valid samplings across enough sites in enough years for meaningful scientific analysis and answers. On the other hand, I can also testify to how gratifying it is to contribute useful information that helps conserve rare butterflies. Most amazing and joyous is documenting specialist populations thriving against the odds long-term, even expanding and increasing.

INVENTORY

The simplest form of surveying to inventory a site is informally noting what's present and where in a site. More visits in more times of the growing season and in more years will accumulate a more complete list of butterflies. It usually takes several years of visits throughout the growing season to come up with a near complete list. If you keep adding onto the same list over many years, though, this list becomes a well documented historical record but not necessarily an accurate reflection of what is there now. To address this, the inventory list can be "refreshed" (started over) periodically. If you do not do this, you may be carrying over species found a long time ago and never since.

By periodically starting a site list over, inventorying can become a kind of informal monitoring. This process can chart how many species, and which species, were recorded throughout, or lost or gained during the various inventory periods. However, I'd like to caution that monitoring (assessing the occurrence or abundance of a species over time) usually requires accounting for effort. Otherwise, apparent gains and losses can in fact be due to greater or lesser survey effort, rather than due to an actual change in the butterfly fauna. For example, a species may have been mostly or entirely overlooked in the site in the earlier period due to little or no survey effort in the times and places necessary to find it. It may not even be possible to figure out whether survey effort occurred in those times and places or not. In the later period, with many enthusiastic dedicated surveyors, the species may be "discovered" at the site. It's not possible to know whether the species has been present all along, and possibly even declining from higher abundance in the earlier than later period.

Some research has demonstrated that the most efficient way for inventory visits to accumulate the most species on a list the quickest is for experienced observer(s) to "wander" (go where they think they should to find the most species), rather than follow a set route. That's because experienced observers are familiar with the fauna in range of the site and have search images for the kinds of places where particular species occur and where to find the most butterflies (e.g., productive nectar flowers). Unless set survey routes systematically cover the entire site in thorough coverage, these routes may not take the surveyor by these particular microsites on a given day. This wandering does not have to be an informal exercise. For each date, you can still keep track of how many individuals you saw of what species, and *how much time* you spent doing this (and distance covered, if it's possible to measure that too) as well as time of day. You can also keep track of your results by subsite (unit) of the site (e.g., woods, old field, marsh, etc.), and record weather information. The more information you track, the more you can evaluate how much effort you spent searching in a particular species' flight period and habitat. This makes it possible for your project to be both an inventory and monitoring. But there is a tradeoff that the more effort you spend recording information about your surveys, the less time available to look for butterflies. Likewise, you may not feel able to process all that additional information in the compilation stage. Then it may feel as though the effort was wasted. On the other hand, I advise that if it's possible to record the data, do it, and some time later it may become possible to process it, perhaps by sharing it with someone else.

Total species lists or species-by-species compilation? Often an aspect of surveying is the compilation of a single list totaling the number of species found ("species richness"). How many different kinds of butterfly species did you find in a site? This is an interesting statistic, but it has the effect of reducing all observations to presence-absence with abundant residents and rare vagrants accorded equal weight in that single number of how many species were found. Unless the species are subdivided into ecologically meaningful groups (such as prairie specialists, migratory generalists, and so on), different sites can be similar in richness but still have very different faunas relative to conservation value. Or the same site can hold steady over time in species richness, yet lose its local specialties while gaining some common or vagrant species. That's why I encourage you to pay attention to "species composition"—that is, which species are making up those totals. It is helpful to classify species into affinity groups, to see whether different groups have different patterns of response. Degree of specialization is just one way to categorize. For example, some analyses group butterflies and moths by what life stage they overwinter as, or where in the vegetation they consume food as caterpillars (in the grass layer, or shrubs, or treetops), or how many kinds of food plants they are known to consume (one, few, or many).

MONITORING

While listing is useful, especially on large scales, I consider the first level of formal monitoring to consist of recording specific dates, locations, and survey effort as well as species seen, preferably the number of each of them. You can work up from that first level of monitoring in graduating levels of increased effort, number of sites, consistency, and detail. Are you able to identify butterflies, or recruit staff or volunteers to do it for you? Are you identifying all butterflies seen, or targeting only one or a specific set of species? Your survey dataset is more useful the more it is possible to identify exactly which set of species you were surveying. The more informal the effort, the more participation you may obtain, but the less analyzable the data are scientifically. The more "rigor" (rules and requirements) you bring to the survey protocol, the more scientifically analyzable the results are, but this usually comes at the cost of sample size—fewer participants and fewer surveys.

Since you need to master the common look-alikes if you are targeting specific rare species, it makes sense to record at a minimum all those common look-alikes when conducting surveys specifically targeting rare species. This not only provides a way to evaluate how well the surveyor is doing at distinguishing the rare from the common species, but also gives you more data to work with. I've been amazed how much these common species have taught me and helped me understand rare ones! Some phenomena apply to butterflies generally, such as abundance fluctuations and seasonal shifts, while specialists may be more sensitive to other factors, such as habitat management. On the other hand, more frequently recorded species closely related to but not as specialized as the rare species may register a pattern not yet apparent in the rare species' data but worth deliberately trying to study in the rare species.

To maintain consistency throughout the monitoring program, it is important to avoid changing method during the program. You need to try to anticipate what is feasible so as to pick something that can be executed consistently throughout. As long as the method picked is appropriate, it is better to stick with that than switch to something else that may become more fashionable. If necessary, do a pilot project for a year or two to test out different methods, so that when you go for it full fledged, you've got the method set for sure.

A pilot program can also help you avoid being too fancy about your method. I've seen more than once that an agency or committee decides to take on butterfly monitoring but have little personal experience yet with doing it. Their protocol often looks "fancier than thou" to me because they specify exacting protocols for weather and time of day and minimum full week intervals between all surveys to reduce the chance of counting the same butterfly individuals on more than one survey. I understand the laudable desire to control these variables as much as possible. But then the difficulties of execution set in. Midwestern weather patterns do not accommodate short butterfly flight periods and other surveyors' commitments. Not allowing the next survey to be 4-6 days later can mean that the next staffable survey date within protocol is 12-14 or more days later, missing the peak flight.

Butterfly surveying is "fuzzy." I advise against fighting this fuzziness by having ever more exacting protocols that mean way fewer surveys actually happen. Instead, I recommend that you embrace this fuzziness by doing lots of surveys to even out those inescapable factors. At the same time, measure those unavoidable variables of weather (amount of sunshine, temperature, wind speed), date, and time of day, so that you can learn more about what is and isn't affecting detection of particular butterfly species.

If you want to make a change to your survey method part way in, some techniques can improve comparability within your program. First, you can add instead of change. That is, continue the existing program in some form, perhaps reduced to accommodate the addition of another survey method. This maintains some long-term continuity while also obtaining the benefit of the new improved survey program. Second, run the old and new survey programs simultaneously for several years, to calibrate them to each other.

Fixed survey routes are particularly appropriate when you want to assess abundance of butterflies over time. Keeping the survey areas the same controls the variable of location. You have to do some scouting work first, or have experience already at identifying appropriate areas to survey, because you want to place routes appropriately and avoid moving them later. If you find that you do have areas that need to be surveyed but were omitted from the original route, I advise adding them, rather than changing the existing route. You can keep the data separate on the original and added routes, and that way maintain long-term comparability while also obtaining the additional information.

Surveying along a route is often called "transect surveying" because the survey is conducted along a route corridor, or "transect." You need to decide how wide your corridor is. The gold standard of butterfly monitoring, The British Butterfly Monitoring Scheme, uses a strip that extends 2.5 meters to either side of the surveyor—in other words, a 5 meter strip. Surveyors record only those butterflies seen within 5 meters in front of them or 2.5 meters to either side. Some prairie researchers use a strip 10 meters wide. Scott and I use an unlimited width strip, adapted from some kinds of ornithological surveys. So long as we can identify the butterfly and keep track of it to avoid double-counting, we record it. We freely recognize that the effective width of our strip varies by species. Regal Fritillaries are identifiable at a considerable distance beyond 10 meters. I don't want to restrict my survey to the few that get to within the bounds of a set route corridor. On the other hand, most Ottoe Skippers have been much nearer to me when detected and identified than the bounds of the British strip. An unlimited strip means the observation rates of different species aren't directly comparable because we are recording Regals over a larger area than Ottoes. The narrower the strip, the more such direct comparisons among species are valid, but at the costs of having to spend time deciding whether a butterfly was in the strip or not, and not recording some butterflies because they were only seen outside the strip.

Straight line routes or contour routes? Scott and I use straight lines where the habitat is linear (roadsides and powerline rights-of-way). We also use rectilinear routes where the habitat is large, level, and relatively uniform. Such routes are easier to replicate than curvilinear ones. However, where the site has topographic relief, I prefer contour routes that maximize staying at about the same elevation (in the same vegetative type) as long as possible, then making a beeline for the next unit (a different vegetative type), to maximize sampling within a unit while also minimizing edge effects along the boundary between units. Some people like to set out straight parallel lines regardless of topography and management history because this systematically covers the site. There is merit in that. It takes the subjectivity out of route layout, avoiding the risk of discrimination. However, it comes at the cost of making it more difficult to group data by vegetative and management characteristics. For understanding what vegetative types are more or less used by the species, and to track management preferences, I find it much more useful to group survey results at the scale of units that can be classified by vegetative and management characteristics. It would also be possible to do this with straight line routes but would be more cumbersome as the butterfly locations would have to be individually cross-referenced to these site characteristics. A variant of parallel routes is random placement of segments, for the purpose of gaining statistical advantage for generalizing. But it comes at the cost of inefficiency, because of the down time moving between one segment and the next, rather than surveying the entire time, and risks missing places that you know are valuable for the species you want to study.

Although European monitoring programs pioneered the protocol of weekly surveys at the same sites during the entire growing season, most of us in North America have not done this. In the cost-benefit analysis, weekly surveying usually only works in well known faunas in areas of high human density. That's because it's usual practice in lesser known areas and faunas to try to collect more basic data on what lives where in more sites, rather than obtain more detailed data on fewer species at many fewer sites. Most of those European sites surveyed weekly were selected *because* it was already known those sites were most worth surveying for butterfly monitoring. In regions with relatively few personnel available for surveys, it's usual practice for them to field themselves in more targeted ways, focusing on particular species or vegetation type, rather than limiting themselves to more thorough coverage at fewer sites. Even where weekly surveys occur season-long, random sampling is rare. It's hard to get the motivation to sample the dud sites that keep turning up in such schemes. Most sampling is focused on sites of pre-existing interest because they are preserves or known sites for certain butterflies or vegetation types.

One alternative is the "reduced effort" monitoring scheme, which I prefer to call "targeted effort" because it's possible to put a lot of effort into these surveys. In this method, fixed routes are established but surveyed only during the flight period of the targeted butterfly. All species can still be counted, of course, and in doing so, butterflies that have similar flight periods can also be effectively surveyed. Much of our butterfly field work moves from one target species to the next—from Frosted Elfin (and Olympia Marble, Gorgone Checkerspot, and Cobweb Skipper) to spring 'Karner' Melissa Blue (and Dusted Skipper and Phlox Moth) to Regal Fritillary (and Gray Copper, Ottoe Skipper, and Leadplant Moth) to summer Karners (and Gorgones). As Scott and I move through this rotation, we also get effective coverage of Brown and Hoary Elfins, Mottled and Persius Duskywings, and Great Spangled and Aphrodite Fritillaries, to mention just a few of the other butterflies adequately covered by the places and times we field ourselves for specifically targeted butterfly species. Our goal is one effective survey per site during the target flight period, or two surveys (usually 1-2 weeks apart), depending on the species.

Season-long coverage can also occur but at less than weekly frequency per site or area. In our surveys of bogs (the subject of another article published online by SWBA), we rotate across regions of northern Wisconsin, which is a rather large place. We can sometimes double up two regions in a weekend, but others weekends we may get nothing done there due to poor weather or other research commitments. Thus, we return to a given area every two to three weeks on average from spring to late summer. We try to finesse the phenology, which varies a bit among regions, by targeting a particular area when we think the seasonal progression is optimal for the target butterfly. But there are limits to that—some weeks are simply rained (or even snowed) out! As a result, we miss main flight for a butterfly species in a given region in some years. But this approach can still work fairly well, if in analysis you're willing to throw out some sites in some years for some species because the surveys weren't effectively executed and adequately comparable. Again, the tradeoff is how much rigor you want compared to how much volume of data you want. Other studies have found that five or six visits 2-3 weeks apart can be fairly effective at finding most species occurring during that season in that place. However, species with low numbers and/or short flight periods may not be effectively surveyed without careful targeting of their flight period.

PRINCIPLES OF TRADEOFFS

One site or many sites? If you concentrate your effort on one or a few sites, you can obtain more data on each of them. However, it's only by surveying many sites that you can determine what's a regional pattern (such as annual fluctuations due to climatic variation) compared to site-specific issues, such as vegetative changes, management treatments, and so on. I can't tell you how many times I've thought something I noticed at one or a few sites would generalize more broadly (e.g., the species only lives in this kind of vegetation), only to find that my sample of 10-20 sites for that species would not support the belief suggested by what turned out to be an anecdote. As a result, I strongly encourage you to allocate your effort first to covering at least 5-10 separate sites for a species, before precluding more sites by lengthening routes in the ones you've selected. It's been my observation that most people's intuition is calibrated too conservatively when evaluating the tradeoff of whether to linger or move on to a new site. That is, we tend to think we should hang on longer here, when statistically it's more beneficial to move on to surveying a new site. At most sites, 30-90 minutes is plenty of time to survey the area's butterflies on that day, after which it's time to move on. If sites are spaced 1-2 hours of travel apart, perhaps the cost-benefit analysis might suggest staying on the longer end of that time range per site. If sites are clustered closely, it becomes harder to justify long surveys at one site if this reduces the number of sites covered. Of course, someone else may find something at a given site that day that you missed. But your monitoring program is immensely strengthened by covering as many sites as possible that you can allocate an adequate effort at.

Measuring effort. This can be done in terms of observation-days for informal methods. It can be in terms of hours or miles walked for more formal methods. Or the method can have a set amount of time or distance per survey, which standardizes effort among surveys.

Random vs. targeted sites. Statisticians love randomness, or its step-sibling representativeness, because this makes it possible to generalize beyond the survey sites. A random sample looks at all possible sites in a region and randomly picks survey locations from that, like a lottery. A representative sample identifies each kind of site desired in the program and ensures that an analyzable sample of typical examples of each is being surveyed—e.g., 8 dry prairies and 8 fen wetlands and 8 pine barrens. However, you need an overwhelming sample of random sites for rare species to register in the survey program. This can be true even for representative sites. Thus, a survey program can only generalize to the extent that species are effectively sampled. As a result, we do targeted sampling. We try to get a large enough sample of sites where the rare species can actually get found to track those populations over time to achieve monitoring consistency among years. That way we have some independence (multiple sites that can be generalizable to the extent that they represent typical habitats and populations for that species) but we also have any sample at all on the butterfly to work with! For really rare species, one can attempt to survey every one of its known populations in a certain geographic area, which diminishes the statistical advantage that random or representative sampling has. That advantage is not eliminated, however. Those known sites may not represent the full range of vegetative and management conditions that species inhabits. That's why it's useful to make some effort to explore outside what's currently known and give unexpected situations a chance to be discovered.

Geographic scale. A lot of work gets done on the scale of the site. However, for understanding how butterfly populations function in a site, I find the unit (subsite) more useful. A unit is a vegetative area relatively uniform in vegetation type, degree of degradation, and management (both type and year of last management treatment). The tradeoff is that the more uniform a unit is, the smaller it likely is. These units need to be large enough so that they do not exaggerate the sampling error of not finding a particular butterfly species that is actually in the unit. In other words, a unit that is 10 meters and 30 seconds long produces an exaggerated abundance rate if any butterfly is seen but most butterflies are going to register as zero even though they are in that unit. I'm throwing out an approximate *minimum* route length per unit of about five minutes and a tenth of a mile (about 160 meters) but preferably ten or more minutes and about a quarter mile (about 400 meters). That assumes multiple units per site. I'd also want at least 15 minutes minimum at a site, preferably 30, with more at a larger site. but then we typically break very large sites down into multiple distinct sampling sites.

Frequency of surveying. The more informal the effort, the more frequently it can be conveniently done and manageably digitized. Listing (checklisting) can even be done daily. Scott and I try to record all our informal butterfly observations in a field book (paper and digital). Even butterfly survey routes can be done quite frequently. The gold standard of butterfly monitoring is the weekly butterfly survey conducted season-long, and that certainly seems as frequent as surveys need to be done. Instead of that, many survey programs target a specific time of year or species or set of species.

Both inventory (presence-absence or species lists) and abundance monitoring are prone to the same issues of detectability and effort. Abundance data can appear more difficult to make comparable among surveys, since the number of a butterfly species found depends on a number of factors. How can all these be adequately controlled? Even under a wide variety of conditions, the species may still be findable, even if in quite varying numbers. We butterfliers experience this a lot, when we visit a site early in the morning or in poor weather or early in the flight period compared to visiting again later in the day or in better weather or later in the week. However, whether you find any individuals at all is just as dependent on the same factors as whether you find many or few. So the problem of "false negatives" (finding zero when the animal is actually present) is just as much a problem in presence-absence analysis. But the full range of possible positive occurrences (from occasional vagrant to abundant resident) is compressed into a single value (present). As a result, presence-absence is a weak way to detect a pattern. In an abundance analysis, you can detect a decline of a locally abundant species when it goes from abundant to just common, or from that to just reliably present in low numbers. In presence-absence analysis, the change only registers once you can't find any at all and have reliably distinguished this observed absence from a false negative. Whether you are examining the species' trend in one site or for a region, either way you want to know about a problem when there's still time to intervene, not once the species is absent.

Presence-absence data can be directly compared only if effort at all the sites is similar. Otherwise, there are statistical approaches to try to correct for effects of effort. But these are beyond my scope to describe here.

With abundance data, observation rates can be used to make data comparable among surveys of varying length. As long as a reasonable minimal effort is conducted per site, so that zero reflects either low numbers or absence (something I call subdetectability) rather than inadequate effort to find what is actually prevalent, observation rates (individuals per time or distance) can account for unequal effort per site. It's important to keep an eye on both observation rate and the area this occurs in. A higher rate in a very small site may be a smaller population in absolute numbers than a lower rate over a much larger area.

COMPARABILITY

Within a site, population monitoring becomes more valid as more attention is paid to making individual surveys comparable (consistent) among each other. Comparability issues arise out of the survey method (who is doing the surveying and how), date (flight period), and weather.

My first priority for striving to make surveys consistent is experience of the surveyor. If the same surveyor(s) do all surveys and are already experienced at the start, then there's little variation in surveyor skill throughout the monitoring program. If different surveyors do the work on different days, this variable can still be comparable among all surveys if the surveyors are all experienced and qualified at the start. Many studies have found similar results across different surveyors under these conditions. However, if there's large variation in observer skill (either among surveyors or for the same surveyor starting as a novice and becoming an expert), that could affect survey results a lot.

Next is how wide is your survey area? Are you counting individuals no matter how near or far, so long as you can accurately identify them and adequately avoid double-counting them later? Or are you only counting those seen within a certain distance of you? The British survey 2.5 meters to either side of the surveyor. Some prairie researchers cover 5 meters to either side. Scott and I use the unlimited method, as many bird surveys are done. It is possible to calibrate results among different strip widths, when comparing one program to another. But strip width needs to be consistent for all surveys within a monitoring program.

Next is how fast you are moving. Do you count all effort from start time to end time, regardless of whether you go slowly and take lots of pictures on one day, or zip around the next? Or do you try to make the effort similar among days by going at about the same pace and deducting out timeouts for rest and photography?

Seasonal timing is also critical. The surveys all have to be in the main flight period of the butterfly in question. Surveys outside that are much less effective for portraying the butterfly's occurrence and abundance.

Weather falls further down my list of concerns. That's not because weather doesn't matter. It's because it takes more judgment than you might expect to evaluate how weather is or isn't affecting the results. Some of our highest counts for target species have occurred in remarkably mediocre weather. I'd get more rest during the butterfly field season if only this weren't the case! Especially in droughts, it has appeared to me that some butterflies may become less active when it is warm and sunny (perfect butterfly weather) because desiccation is a bigger issue for them than calorie expenditure to be active when it is cooler and overcast and not as dry. As a result, I use the method of positive evidence for dealing with subpar weather. If butterflies are active and detectable, then it's suitable for butterfly surveying. If the weather is "butterfly weather," then it is an analyzable survey whether good butterfly results occur or not. In the weather tradeoff, I definitely bias toward doing more surveying and then seeing how much of it I can make a case for including in a given analysis. And we definitely do try to resurvey if the only weather available so far has been mediocre. That's how I've noticed that weather doesn't always work out the stereotypical way, although often it does.

If you're looking toward getting the best surveys in the least effort, then you may set your weather standards higher than I do—just beware that you're not so picky that the main flight period is over before you can get your surveys in! The more stringent you are about weather, the less surveying you get done. The more relaxed you are about weather, the more you have to use judgment, and defend that judgment, in determining what should and should not be usable surveys. The more you discipline yourself to use objective criteria to apply across the board to include and exclude surveys in analysis, the more scientifically defensible your findings should be—for example, all surveys after a certain time of day, or above a certain temperature, or meeting a combination of minimum temperature and maximum cloudiness.

With a broad protocol, you can analyze individual species to determine which factors are more important in affecting detection. These factors may even vary from year to year. In extreme drought, moisture may be the primary activity limiter. In a cloudy wet year, sunshine and warmth may be the primary limiter. Sometimes surveys outside the standard protocol for time of day and acceptable butterfly weather yield higher numbers for certain species than surveys on the same date that are mid-day, warm, and sunny. For small butterflies, and ones that can't be identified in flight, surveys may not be as successful when it's mid-day or afternoon on a warm sunny day because of how active the butterflies are. They may be so frenetic as to be difficult to spot, much less track, much less observe closely on a perch. Alternatively, in very hot or very dry weather, some species may adopt heat-minimizing behaviors, becoming less active and perching in shaded (hidden) locations. Meanwhile, of course, other species may be going berserk, occupying all your attention. Just as with birds, so also with butterflies, it is not possible to devise one survey protocol for time of day and weather that is optimal for all species present in a given area on a given date. However, as also for birds, it is possible to come up with a protocol that is broadly acceptable for most species, and develop additional survey programs to cover the remaining species effectively.

Even though I prefer fixed routes, or similar routes as much as possible among surveys at the same site, doing that actually falls pretty far down my list in comparability. That's because analyses in prairies show that different experienced observers, following different routes in the same set of sites, came up with well correlated results compared among sites. When the locations were matched on the scale of the unit (portion of a site), the correlation reduced but was still significant and positive. Matched at the cruder scale of "site," but counting all effort at the site (which was more effort than in the unit analysis), the stronger the positive correlation in survey results. In this case, though, all surveyors were targeting the same prairie specialist species to focus on (while counting all species found), and so were interested in the same habitat (vegetation) types. Even though the routes weren't the same, the habitat was. So I advise organizing surveys by vegetation type—that is, set survey routes so that data are kept separate for different vegetation types.

Variation in amount of obscuring vegetation among sites might cause differential detectability of the same butterfly species. The concern is that brush and trees in some sites reduce detectability of a species compared to surveying the same species in wide open habitats. I would add that level sparse grasslands appear easier to detect butterflies in than rocky steep slopes or tall thick grass or weeds, not just because there's less vegetation to obscure butterflies but also because surveyors need to spend less time watching their footing. A method used to address this is line transect extrapolation, but this has issues of assumptions and execution (see the Extrapolation section). A simpler approach is to group sites by vegetative type and structure and make comparisons among sites within but not among groups, or with cautious interpretation of comparisons among groups. I am interested in more research to evaluate how much of an issue this is.

IMMATURE STAGES

Surveys need not be limited to the adult life stage (butterflies). One or more of the other three life stages (egg, caterpillar, chrysalis) can be surveyed also or instead.

Immature stages are usually more time consuming to find, but not always. They are smaller, less mobile, and usually better camouflaged (less noticeable) compared to flying and sometimes brightly colored butterflies. Immatures may be obscured, compared to adults up in view nectaring or flying or defending a territory. Some caterpillars are active more at night. Plus, identification materials are not as advanced and comprehensive for immature stages.

But in some cases, immature stages are easier to survey. Examples especially apply to caterpillars, followed by eggs. In the case of these caterpillar examples, they are readily detected because they feed colonially, or because they do not hide when feeding on a particular plant species that is easily searched, or because they make distinctive, readily noticed feeding "signs" (often called damage) or shelters on their food plant. Sometimes the insect itself isn't surveyed directly, but just the extent and abundance of those feeding signs. The 'Karner' Melissa Blue illustrates both these approaches. The caterpillars are distinctive, active during the day once it's above about 52 degrees Fahrenheit, and the ants that tend them (beware getting bitten) just make the caterpillars easier to detect. However, these caterpillars make distinctive feeding signs on their caterpillar food plant (wild lupine), which can be assessed even more quickly. Likewise, eggs (such as those of the Monarch) are readily identified and findable on milkweeds. Scott and I hosted a field trip with Dr. Robert Michael Pyle to search for Karner eggs in November in central Wisconsin in preparation for his book *Mariposa Road: The First Butterfly Big Year* (2010, Houghton Mifflin Harcourt, Boston). We each found one of these distinctively shaped eggs in dark cold winter weather when there was no hope of surveying effectively for any adult butterflies. As a result, immatures may be surveyable at times of year and in weather conditions unsuitable for adult butterfly surveys.

An important advantage of surveying immatures is that breeding habitat gets directly pinpointed. The caterpillar life stage has typically been found to be the limiting life stage for rare and declining butterflies. So identifying the important areas for those caterpillars, and what it is about those areas that is so important to those caterpillars, is very valuable for conserving that butterfly species.

The disadvantage of surveying immatures is that these are usually single-species surveys not practical to do on a large scale for multiple species. Those monitoring programs I know of that target immatures focus on one species, or a set of species using the same food plant. It's hard for me to imagine how it would be possible to survey immatures on the scale and speed that most adult butterflies can be. As a result, so far it appears that immature surveys are a more specialized survey method used to complement (not replace) adult surveys.

Usually caterpillar counts correlate well with adult counts done in the same area. That is, when a survey program surveys abundance of caterpillars in a number of sites, and then the subsequent flight of adults, the caterpillar and adult butterfly counts correlate well. As a result, this validates both kinds of surveys. Although adult butterflies do fly around and disperse out from where they grew up, their concentration areas usually correspond to breeding habitat. The primary exceptions are when the butterflies are migrating (e.g., Painted Ladies streaming through my front yard or a mountain pass) or hilltopping (an interesting phenomenon I've seen on prairie ridges in Wisconsin, thought to be analogous to the "lek" behavior of birds that congregate to do breeding displays). So adult counts appear generally effective at identifying the vicinities important for supporting resident butterfly populations. On the other hand, caterpillar counts are also validated as accomplishing that as well.

EXTRAPOLATIONS

Everything I've discussed so far assumes data compilation based on numbers found, but there are also several extrapolational methods. These are usually used as a method of inventory rather than a long-term monitoring method due to how labor intensive they are. The extrapolation techniques try to calculate how many individual butterflies really were there, including ones not actually seen, based on the distribution of individuals seen relative to how far away they are from the surveyor, or based on how many individuals marked from a previous day are among the individuals found today. These ratios are used to infer how many individuals were actually living in the habitat compared to how many found. There are lots of assumptions required for these methods, and difficulties meeting these assumptions.

One long-used extrapolational technique is mark-release-recapture (MRR), where each individual that can be caught is marked and released. The next day, some individuals seen are marked ones (recaptures) while others are new (not previously marked) that are then caught, if possible, to be marked. If enough individuals are found over enough days, the ratio of recaptures to newbies can be used to estimate how many individuals were actually present. Unfortunately, it's been established with a number of butterfly species that if the butterfly is active when it's being handled (instead of being marked when "asleep" on a roost or when anesthetized or otherwise unaware), it becomes much harder to recapture compared to ones that weren't aware of being marked. In other words, the experience of being handled is aversive for butterflies. Since it can be hard to mark unaware butterflies, that is rarely done. This isn't a minor difference: e.g., less than 10% recapture of butterflies aware of the handling compared to greater than 80% recapture of individuals that weren't aware. As a result, most MRR studies have very low recapture rates. This has the mathematical consequence of producing much higher estimates of undetected butterflies assumed to be present compared to when individuals of the same butterfly species are undisturbed by the marking technique and remain resident and readily re-findable in the place where first seen. This has the conservation consequence of assuming more butterflies exist in the site when recapture rates are low.

Another technique is line transect (LT), where the distance perpendicularly off the survey route each butterfly is seen is used to estimate the "decay" rate of detection based on more butterflies seen nearer and fewer farther away. This method starts with the assumption that you find all butterflies present on the pathway you are traversing, to establish the baseline for 100% detection. However, when Scott and I watch each other surveying, we see that in fact some butterflies perched between our feet do not flush, or flush out behind us, or flush to the opposite side of where we happen to be looking, and so do not get detected by the one who almost stepped on it. More importantly, this method is more suited to surveying large animals by plane, boat, truck, or other vehicle, with one person operating the vehicle and one surveyor on each side looking out the side windows. When conducted by one surveyor walking a butterfly route, there's a strong bias to look ahead, so as not to get injured! This biases butterfly detection to be in a narrow strip near the surveyor. For example, most Red-disked Alpines I see are straight ahead of me, because that's where I have to look so as not to get my eye poked out or leg broken as I traverse the bog. However, these alpines are often 10-20 feet or more ahead of me, and may have lots of intervening shrubbery in the way. It's true that butterflies become more difficult to detect both the farther away they are and the more intervening vegetation is in the way. The assumption of LT is that the farther off line a butterfly individual is recorded, the less that vegetation is confounding or obscuring its detection. But in the case of the alpine, I would only be measuring distance off my survey route for LT, not how far away from *me* the alpine was. In the LT equation, it appears that only alpines I practically step on get detected, but believe me, a lot of the time, I'm never nearly that close! It's also difficult to estimate where the butterfly was when first seen if it was flying because you are tracking and identifying the butterfly first, not flagging its original location. But this distance is critical to valid execution of the extrapolation equation.

These extrapolational techniques are also subject to all the detection issues that affect regular transect surveying. If it's better weather or better timing in the flight period, more butterflies get found and therefore more get assumed to exist but not be found too. These techniques do not have a way to get around this issue, because they use the patterns in what got found to assume what didn't get found. For example, if extrapolations differentiate between male and female, the estimates are usually much higher for males than females, because males are usually more detectable than females and so the extrapolation equation gets more data on males to work with. Furthermore, extrapolational methods assume undetected butterflies behave the same as detected ones. But we know that's not entirely true. Behavior is a factor that contributes to which individuals get detected—they behaved in a way that made them relatively more detectable than undetected ones.

Extrapolation is rarely used for long-term monitoring because it is too labor-intensive. MRR requires a lot of time spent on capturing and marking individuals, and LT a lot of time recording distances, which reduces time available to look for butterflies. An additional downside of MRR is the risk of injury and mortality to the butterflies during capture and handling. As a result, I have not seen MRR used to generate monitoring datasets for each year over a number of years at a number of sites for individual target species, much less multiple species simultaneously. When LT is used, often the distances are measured only in an initial phase to calculate the ratio of number observed to total assumed present. From that, the ratio is used to calculate a constant used as a multiplier on subsequent survey results to extrapolate total present.

I find it difficult to relate extrapolated results to what gets seen on transect surveys. Extrapolational techniques generate estimates that are usually one or several orders of magnitude higher than the actual number of butterflies seen. It's hard for me to relate those extrapolated results that are in the hundreds or thousands to what it was actually like to be present in the site when the butterflies were being surveyed. Transect survey results can be expressed in ways that are concrete (number seen per time or distance, or over a given area). It's easy for me to visualize what someone else's transect survey results were like to experience. To promote comparability among different studies and programs, extrapolational surveys could express results not just as extrapolations but also as observation rates, as transect surveys do.

In tests of similarity of outcome, transect survey covary strongly with MRR and LT extrapolations. Thus, my biases about methodology are relatively minor compared to the major differences that butterfly populations exhibit within site among years (fluctuations) and among sites (small or large populations). Dr. Jeremy Thomas expressed this another way: relatively rough methods are sufficient for ranking sites by population size (large or small). I would add to that the ranking of fluctuations (good or bad years). Transect surveys have also been "calibrated" to an extrapolational method by doing both simultaneously in the same areas. Then in future surveys, butterfly counts are multiplied by a constant to estimate extrapolated population size.

Once minimal basic requirements are met (measure of effort, adequate sampling method to detect butterflies), picking a survey method comes down to tradeoffs. Extrapolational methods are much more labor intensive. As a result, fewer species and sites typically get covered for fewer years with these methods. As a matter of practicality, large-scale long-term multi-species butterfly monitoring programs have usually used versions of transect surveys.

I emphasize consistency within each survey program, but I happily embrace diversity among programs. The actual numbers generated by different methods are not directly comparable. That is, 8 from a transect count isn't the same as 8 from an extrapolation. However, the changes up and down within one program can be compared to the ups and downs within another. For example, the North American Butterfly Association's 4th of July Count Program can be a useful source of information on regional fluctuations, if enough counts occur in a region and they get adequate samples of the species of interest to you. You can use that to provide context for a more rigorous monitoring program at your site(s). Different programs can complement each other, allowing both broad-based regional sampling and focused detailed surveys at a few sites of particular interest and conservation need.

PLANT MONITORING

It is highly useful to include some plant and vegetation surveying in a butterfly survey program. Assessing distribution and abundance of caterpillar food plants and nectar flowers can be very useful for understanding butterfly results. Classifying survey areas by vegetative characteristics (type, degree of degradation, management type, years since last treatment) is also very useful for identifying which factors are more or less important in explaining why butterflies occur where and in what abundance. This plant and vegetation monitoring needs to be efficient (not time consuming) so as not to reduce the size of the butterfly dataset being obtained. The plant monitoring also needs to be prioritized to address issues of most apparent relevance to butterflies. The methods described here are rough, intended to do general classifications that can get a lot data "bang" for the effort "buck."

A simple categorical approach is widely used for assessing abundance of plants. The British use these categories: Dominant, Abundant, Frequent, Occasional, and Rare (abbreviated as DAFOR). Scott and I adapted ours from guidance provided by the Wisconsin Department of Natural Resources for 'Karner' Melissa Blue surveys. We use these names: Superabundant, Abundant, Common, Uncommon, Rare, and Zero detected (or SACURZ for short). We define these categories on approximate order of magnitude steps. That is, rare means about 1 stem or clump in a given area, uncommon about 10, common about 100, abundant about 1000, and superabundant about 10,000. We also allow intermediates between each of these: that is, rare/uncommon means about 5. As we walk through our survey area, we classify the overall abundance of a plant species using these categories. For several years, we used this approach to assess nectar abundance by flower species. Since then, we have been more limited in our plant assessments to key caterpillar food plants.

This categorical approach has the advantage of requiring relatively little time to do in the field but the challenge is to make this subjective or informal approach as consistent as possible among surveys and sites. You can improve comparability by providing some specifics of how you define these categories, or photographic examples of each category. It's also important to be clear whether you are assessing flower abundance or plant volume (amount of plant growth evident independent of whether it is flowering or not), or both. In addition, it is useful to note when you think your assessment is not a fair reflection of actual abundance because it is not the correct seasonal timing to do so.

Sward (turf) height can be assessed by the "yardstick" method. The idea is to measure the height of the preponderance of the vegetation, excluding the few tallest plants and flowers. This requires an overall visual assessment of the unit, but then taking a single measurement, leading to time efficiency.

Canopy (shrub and tree) cover can also be approximated. One approach we use is categorical: prairie (0-10% canopy cover), open savanna (10-25% cover), closed savanna (25-50% cover), and woodland (50% or more). Another approach we use is to assign an approximate shrub cover percentage to a unit (e.g. 5%, 40%).

Volume of dead plant litter on the ground can also be measured categorically. We have used this for grassland bird study, but I have seen this as a variable relevant in butterfly analyses so I hope someday to look at this in our data as well. We started with these five litter categories: none, light, moderate, heavy, and very heavy. We added in the intermediates between these categories (none to light, light to moderate, etc.), amounting to nine categories in all. In the last few years, we've encountered some units that have even more litter than our very heavy category, so we've added a tenth category for that: super-extra-heavy litter (!).

Vegetative classifications can be done on the scale of survey unit. The goal is to classify by rough or obvious distinctions: weedy old field, high-quality never-tilled hill prairie. Preserve guides can be very helpful in identifying the kind of vegetation present at those sites. Once you've mastered that, you can then cross-reference those categories to other sites not specifically covered in a guide. Using classifications provided by botanists can be very enlightening in butterfly surveying. Sometimes botanical perspectives relate well to butterfly data—for example, muskeg bogs in northern Wisconsin have a typical butterfly fauna and sedge meadows have their own particular skippers. On the other hand, it can be equally important to notice when botanical perspectives do not line up as well with the butterfly data. For example, even though Gray Copper is restricted in range to mid-continent prairie, this butterfly is not all that well represented in prairie preserves and many Gray Copper sites don't look like stereotypical prairie. It's not that the botanical perspective is wrong. But sometimes I think insects may be useful for helping us see vegetation in a different way, and to understand another side of prairie (the weedy side) that some native prairie butterfly species prefer. It's also useful to categorize degree of degradation as evident by invasion of non-native weeds and brush and amount of native herbaceous diversity. We use three categories for this (high-quality, semi-degraded, and degraded). To the extent you can determine management by direct observation and consultation with owners and managers, it is also very useful to classify your survey area by type (e.g., grazing, burning, mowing, nothing) and when that last treatment occurred. This may not be pinned down to date, and may not be identified to month. But a crude approach can determine whether it happened since the last growing season (less than one year ago), or one year more ago than that, or two, etc.

More detailed systematic approaches to plant sampling are beyond my scope here and in the field when doing butterfly surveys. I encourage a careful cost-benefit assessment when determining how rigorous a botanical survey to conduct. This may not be so much a cost of money (e.g., purchasing equipment) but of time. I raise this point in animal survey methods too. A more rigorous dataset that is small may not be as effective at providing insight as a large dataset with less strict methods. On the other hand, more careful and targeted botanical data collection may be necessary to bring resolution to butterfly questions not yet adequately addressed by available datasets. For that, refer to botanists and the botanical literature for insight into the range of more rigorous plant sampling methods developed. It's also possible that you may find someone else has already done some of the plant data collection for you. They may already have mapped the vegetation or done botanical sampling in your survey areas that you can reference. In that case, you can avoid duplicative effort by applying your effort to something else. Of course, plants also deserve careful surveying and monitoring, as much as butterflies do. It's just that it's beyond the scope of a butterfly monitoring program to do so.

HAZARDS

As enthusiastic as I am about butterfly surveying, there are many natural hazards to this activity. Weather forecasts can be inaccurate, sometimes dramatically so and in dangerous ways, e.g. a thunderstorm developing when we are far from the car. Myriad biting flies, mosquitoes, and ticks can vex as well as transmit disease. Plants can bite back with thorns and irritating chemicals. A rattlesnake may take umbrage at our presence and a mama bear (with cubs hidden in the underbrush) may be reluctant to depart from our desired survey route. Dramatic cliff topography can mean scary footing. Level grassland may obscure treacherous footing around animal burrows.

Of equal concern are human-caused risks. These include familiar ones such as hazards traveling on the road, criminal activity, and so on. In addition, some people find it disconcerting to see someone doing something however benign that they don't understand. As familiar as butterfly and bird surveying is to me, apparently someone crawling around on the ground (in search of caterpillars or phlox moths), or walking around a site with binoculars and clipboard can inspire not just interfering curiosity but also suspicion or intervention by neighbors, passersby, and law enforcement. Be alert to head off these misunderstandings or disruptions.

Please be careful and sensitive to the sensibilities of others!

WHAT I WOULD DO

If I were designing a butterfly monitoring program for a region, I would do a combination of informal and formal, inventory and monitoring. I would aim for three levels of surveying. For my first level, I would want to set up a method of capturing the greatest quantity of informal types of records from as many participants as possible on as large a scale as possible. Published reporting venues and websites designed to collect observations from volunteers are exemplars. Second, I would encourage "wandering" (relatively informal) surveys for exploration of the butterfly fauna. I would especially try to enlist surveys at targeted vegetation types and butterfly species. The goal is to try to find undiscovered sites and track the butterfly fauna in more areas than can be covered in formal monitoring programs. The surveyor defines the site as best he or she can and records start/end time, species seen, and site characteristics as possible. Third, I would establish fixed-route transect surveys targeted to sites and times for finding particular butterfly species that I want to keep a close eye on, but with surveyors recording all individuals of all butterfly species seen. As a result, many non-target butterfly species would also be effectively surveyed. To analyze population monitoring, it is essential to document regional fluctuations. For example, I think of Wisconsin as having at least seven regions: southeast, southwest, east central, west central, and for our very wide northland, northeast, north central, and northwest. I would aim for at least 5-10 independent sites per region per targeted species. Such a program requires an investment in administration to recruit volunteers (especially butterfly enthusiasts), train staff (such as managers) to become more proficient at butterfly identification, and recruit people with experience in program coordination and data compilation. It is important not to start too big too soon but begin with something manageable that allows for the opportunity to grow and expand if that becomes possible.

If I were designing a butterfly monitoring program for my one site, I would scale down my regional plan to that site. I would aim for two levels of surveying. First, I would want to avail myself of the greater abundance of data obtainable from relatively simple surveying volunteered by visitors. No matter how much I think I am covering the site myself, more observations by more people mean more possibility to detect something no single individual will happen to see. Second, equally useful is a set survey route covering as much of the site as possible in regular rigorous surveys. If possible, I would emulate the Europeans by running this route once per week. Otherwise, I would target specific seasonal timings intended to ensure consistent coverage of particular species and maximum species richness (usually early to mid-July in southern Wisconsin and a week or two later in northern Wisconsin). As much as possible, I would want to obtain at least several years of "baseline" surveying before embarking on a significant alteration in the management program.

To become a volunteer surveyor, take advantage of opportunities already available and try to advance in skill level and participation. A beginner can participate in butterfly field trips and training workshops offered by nature centers and butterfly chapters. Many opportunities exist to share butterfly observations in website databases (such as wisconsinbutterflies.org or NABA's Butterflies I've Seen website) and published seasonal observation reports. Participate in a nearby 4th of July Butterfly Count, then start your own and do it in all three count periods. If a butterfly atlasing or monitoring project is occurring in your area, you can participate in that. Or, like us, you can start your own survey program, which we conduct, digitize, analyze, and write up all on our own.

IF ONLY

No matter the level of your interest and skill, I have one wish for your involvement in butterfly surveying. These relate to quality of the data collected, data preservation, measuring effort, and recording the zeroes to prevent misleading positive bias. There's no need to be overwhelmed. Find your place, do that well, and enjoy! Butterflies are wonderful.

If you are a beginner, I wish you would participate in field trips and other programs (many are free) that let you learn from the experts.

If you can identify butterflies, I wish you would record what you see. This could be in a field book or computer diary, or a website database.

If you are recording your butterfly observations, I wish you would record complete species lists (all individuals if possible).

If you are recording complete species lists, I wish you would record effort and conditions (time spent surveying, weather, time of day).

If you record your survey conditions and effort as well as complete butterfly lists, I wish you would have a set of sites you visit regularly to survey the same way each time.

If you do repeated consistent surveys at the same site(s), I wish you would find a way to preserve and share these data for scientific use.

If you are entirely a bystander to butterfly surveying, please express your support for those who are surveying. There are practical ways to do this: donate to support programs, underwrite a study trip, purchase equipment and supplies, provide room and board. But words of encouragement and appreciation can make all the difference too.

Published by the Southern Wisconsin Butterfly Association (SWBA), a chapter of the North American Butterfly Association. To find out about field trips and meetings, please visit http://www.naba.org/chapters/nabawba/