Summary of Science Procedures using Camera Traps

Tavis D. Forrester and William J. McShea
Smithsonian Conservation Biology Institute, Front Royal, VA

The following steps summarize how to apply the scientific method to camera trapping using the same general methods employed by biologists working with the Smithsonian's eMammal project. The goal of this document is to help teachers learn about and share the nature of science during the camera trapping classroom and field activities in the curriculum packet.

1. Generating Scientific Questions

There are two broad categories of scientific questions: theoretical and applied. Conservation biologists (and most non-scientists) are mainly concerned with applied problems. Scientists with eMammal generally use known conservation issues to generate scientific questions. All of these questions must be appropriate with data obtained from camera traps. Camera trap data can provide information on:

- Distribution and variation of animal species/populations in space
- Variation of these species/populations in time

Conventional scientific experiments, with well replicated pairs of treatments and controls, are hard to do in ecology, and with camera traps in particular. As a result, ecologists generally collect data across landscapes or a span of time and use these data to infer a process (described more in **Designing a Study**).

Camera trap surveys can be used to gather baseline data or answer specific questions. A baseline survey might be focused on documenting what mammals are present and in which habitat they are found. Conservation questions addressed by camera trap studies are typically focused on the effects of human activities on mammal populations, from the effects of building a road, to widespread effects from logging, to global climate change. Example questions appropriate to the Virginia-based eMammal projects include:

- What habitat has the most mammal species?
- How are mammals affected by road density?
- Do coyotes and bobcats live in suburban areas?
- What kinds of habitat are associated with high bear detections?

2. Exploring the Data

The first step of any hypothesis generation is exploring the data. This could involve looking at camera images from previous classes, exploring trusted websites for results from other places, or discussing observations made by students. These observations of data, text and stories are the rich ground used to generate interesting questions and multiple hypotheses explaining the reason behind the pattern. The study conducted by the students can have two aspects: Does a pattern exist? (i.e. Are bobcats detected farther from roads than deer?); and what explanations are there for the pattern? (i.e. Deer are avoiding bobcats, or bobcats are avoiding roads, or deer are attracted to vegetation along roads regardless of bobcats, etc.). Often we can only test if a pattern exists and then eliminate 1 or 2 of the many competing hypotheses. The idea is the next scientist (or class) will use your results to advance further until only one hypothesis survives that explains the pattern we observe.

3. Designing a Study

Once a question or hypothesis is generated, we have to place cameras across the landscape in a way that will give us good data to answer that question. This method uses *spatial pattern* to infer an *ecological process*. This requires making assumptions about the process that is occurring and controlling for other factors that might be confusing the results.

An example is the question of bear abundance. We could look at bear abundance across different kinds of habitat to figure out which habitats are associated with lots of bears. The *pattern* is how many bear detections we capture with cameras, and we need to generate questions about the *process* we think is most important. What process is most important for bear abundance? Is it food? Specific types of food, like acorns from oak trees? Is it human development like houses or roads? Places with bear hunting vs. no hunting? Some combination of these factors? We need to determine which hypotheses are the most probable, and which are testable using our tools and resources. Once these are determined, we design our data collection to test that hypothesis.

The first step is to study other scientific findings and generate questions to fill in the knowledge gaps or test if someone else's findings apply to our particular area (i.e. Do studies about black bears in Michigan apply in Virginia?). Then we design a spatial arrangement of camera placements that will test the spatial pattern of bear abundance against what we think is an important process. This is straightforward with one factor and gets more complicated as we add other factors. For example, if we think the amount of oak trees and acorns is driving bear abundance, we can set cameras across a gradient of oak density or in areas with and without oak trees. If we think both oaks/acorns and roads are the important factors, we would sample all four possible conditions: areas with high oaks/high road, high oaks/low roads, low oaks/high roads, and low oaks/low roads.

We just described a study with a simple 1 or 2 factor design. As we increase the number of factors we think are important, the design and statistics needed to test the results become more complex. To test multiple factors together, we might use occupancy analysis, a method to calculate the probability of an animal being in a certain place that is corrected for differences in the probability of detected that animal in the first place. However, this method is beyond the scope of most middle or high school students.

4. Collecting Data

Once we have the study design, we place the cameras on the landscape. The eMammal program recruits volunteers to do this so we can place lots of cameras over a wide area. We still have to follow a design that targets a certain question, allows for realistic time constraints, and accounts for any seasonal differences. Even though volunteers place the cameras and collect the data, is a good idea to get students outside to see local habitats. Classes can use mock camera deployment to learn about local habitats and get a taste for field work.

A large part of data collection is managing, checking, and storing the data in an organized way, but fortunately the eMammal workflow does this all automatically!

5. Analyzing Data

After we have our data, we need to test our original question. The first step (and most fun step!) is to identify the animals that are photographed. These animal detections are the raw camera trap data. The most common ways to analyze these raw data are:

- Relative abundance (the number of detections divided by the number of days the camera trap was in the field)
- Relative Species richness (the number of species at different sites)
- Diversity indices (metrics that are adjusted for survey effort and number of sites; not well measured by camera traps)
- Daily activity patterns
- Occupancy analysis (measuring the probability of a certain species occurring in a site or using a certain area)
- We can measure how different factors/variables are related to this probability
- We can also calculate the probability of different species occurring in the same site to examine possible competition or community associations
- Sampling effort (examining whether sampling captured the majority of mammals that are present)

Relative abundance, species richness, and diversity metrics are built into the eMammal system and are relatively easy to calculate in middle and high school classes. Occupancy analysis is beyond K-12 education.

Usually we have two sets of data: our observation of organisms and our measures of important factors. When we are setting up a project, we think of the graphs we want to create to explain our data at the end. We will place the observation of the organism on the y-axis (it is the response, or dependent, variable) and the measure of our important factor on the x-axis (it is the independent variable). Does the number of animal detections increase with an increase in factor x? We can create a graph and label each axis. Then our data collection is intended to fill the spaces inside the graph. This can be a scatter point or a bar graph, but the intent is to see the relationship between the dependent and the independent variables. Once we graph the pattern, then we can use statistical tests that separate any real differences in the data from the normal variation (a.k.a. "noise") in the environment and study design.

6. Drawing Conclusions

Drawing inferences from data analysis can be straightforward or complicated, depending on the results. It is important to realize that data analysis is observation converted to numbers, where it can be tested and examined (this is why math is useful!). Once tested, it is converted back into a story to explain the pattern observed. The story is created based on the results, but the temptation is the take the story beyond the results. The edges of a known story are where we use *inference* to predict where the story is going. We have to be clear what is known and what is inferred based on our results.

Returning to the bear example, if we found a lot more detections in areas with oaks than in areas without oaks, then we can say that we found a pattern. But what if we found that not all oak forests were the same; bears were detected most in oak forests with a low road density? We might infer that oaks are important but that roads are even more important. Why are roads important? Maybe because lots of hunters go to areas with roads or maybe because roads are related to development and there is less bear habitat in those areas? We can describe the pattern, but we do not yet know the reason for the pattern. Inference can quickly turn into speculation, and once inferences are becoming too separated from data, it is time to move to the **Refining the Questions** stage and generate a follow-up question.

What if we had different numbers of detections but they were not that different? This is why scientists use statistics, so we can test if two values are actually different or are just a result of the random variation present in any natural system.

If the class is using species richness as their analysis tool, they can observe that one area has higher richness than another area, but then they must infer why. Is the richness different due to the habitat types? Disturbance patterns (i.e. fire vs. no fire)? Undisturbed habitat vs. developed? Some inferences may require more investigation.

Caveat: Remember our sampling is always based on imperfect tools place by imperfect humans in a slightly chaotic world. The camera traps cannot sample abundance but only detection (10 bear pictures do not equal 10 bears). The placement of the camera is sometimes too high or too low to detect our species, and it is never exactly the same. The relationship with oaks may be present during normal years but not during years of drought. All of these factors combine to create uncertainty and the uncertainty means you need a large sample size to estimate accurate mean values or correlations between your factors. Because your work is not perfect does not mean you cannot advance our knowledge. You must always review your assumptions and limit your conclusions to fit the data you have and not the data you wish you had.

Refining the Questions – Next Steps

"The most exciting phrase to hear in science, the one that heralds the most discoveries, is not "Eureka!" (I found it!) but "That's funny…" ~Isaac Asimov

It is a truism in science that science answers one question and generates a hundred more. After students have analyzed the data and tested their original questions, they will certainly have more questions. This is a critical part of science and one that should be integrated into the curriculum.

Typical follow-up questions are those that need more data, such as explaining the difference in mammal patterns in the example, or investigating new patterns or ideas that are discovered. An example of a follow-up question that needs more data is the bear and roads example above. Students could check the data to see if there are less bears in areas with lots of roads but few houses, or they could get hunting effort data from the state fish and game website to see if counties or sections with more roads had more hunters.

Seeing a pattern where before there was only chaos is the first step toward new discoveries in science. It is important at this stage to cultivate curiosity in students and get them to look at the data in different ways.

Once the follow-up questions are determined, we start all over again, starting with the knowledge generated by the first class!