Improving Wildlife Surveys with Advanced Genetic and Statistical Tools

Advancements in genetic technologies have provided new wildlife survey tools that are more efficient, less expensive, and allow scientists to detect species that are otherwise difficult to observe. Detection of environmental DNA (eDNA) [ee-dee-en-ay] shed into water or soil by wildlife, is one such tool. Through a multi-year study of Asian carp eDNA in the Chicago Area Waterway System, Dr Martin Schultz of the United States Army Corps of Engineers has been developing new methods that combine genetic techniques with powerful statistical models to estimate the Asian carp eDNA concentrations in waterways.

…

Wildlife naturally shed skin and other tissue along with bodily excretions into their surrounding environments. The DNA contained within these tissues persists in the environment for a short time, allowing researchers to detect the presence of wildlife species without the need for conventional survey techniques, such as trapping and releasing. By using an established technique called polymerase chain reaction, or PCR, researchers can 'amplify' the genetic marker of a target species in an environmental sample containing the DNA of those species present in that environment.

The genetic markers used in PCR function much like barcodes: a unique sequence of the DNA building blocks is shared between individuals of the same species, allowing researchers to test for their presence in a sample of environmental DNA – or so-called eDNA. Results of surveys employing these methods are reported in terms of the number of samples that test positive for a given species, but it is difficult to make inferences about the location and distribution of eDNA sources in the environment with this information alone.

As useful as this method is for detecting enigmatic species or those not typically resident in the sampled habitat, knowing the concentration of eDNA in the samples is crucial for understanding species distributions and making informed environmental management decisions.

Through his multi-year study of Asian carp eDNA monitoring data in the Chicago Area Waterway System, Dr Martin Schultz of the US Army Corps of Engineers has been working to overcome the limitations of these genetic tools. He combined eDNA sampling and PCR with advanced statistical modelling to infer eDNA concentrations from collected water samples.

Dr Schultz could confirm some results of his eDNA concentration modelling by comparing them with reported sightings and surveys of the fish in the area using traditional wildlife survey techniques. At the time the data were collected, between 2009 and 2012, the Asian carp adult population front was located downstream of the Chicago Area Waterway System. He showed that the eDNA concentrations were highest downstream of an electric fish barrier built to prevent fish migration into the Chicago Area Waterway System and that concentrations decreased as one moved upstream into the Chicago Area Water System, toward the Great Lakes.

The Chicago Area Waterway System is being closely monitored for Asian carp because it links the Great Lakes to the Illinois River, where the two Asian carp species of greatest concern, bighead carp and silver carp, are well established. Such close monitoring of bighead carp and silver carp species in

these waterways is imperative to prevent them from reaching the Great Lakes, where they could cause severe ecological and economic damage. Detecting the fish while the population density is very low gives wildlife managers an opportunity to eradicate the fish from the waterway before they become established.

…

Recently, technological advances have enabled researchers to estimate the amount of eDNA in a sample by using quantitative PCR to estimate the number of genetic markers in a tiny subsample of liquid, and then extrapolating this to the whole sample or the environment. However, losses of eDNA during sample collection, storage and processing are difficult to account for and can have a dramatic effect on the accuracy of the estimate. Without a robust and reliable method of accounting for the influence of sampling and analysis techniques on eDNA capture and detection rates, eDNA concentrations obtained in this manner may be rife with introduced bias.

The statistical modelling technique employed by Dr Schultz, called 'Bayesian [bay-zee-uhn] inference', allows him to characterise the uncertainty inherent in each of the model's parameters, and continually update the model to reflect new evidence as it arises. Each iteration of the model after it is updated with new evidence becomes more robust than the previous version.

Dr Schultz used the fraction of water samples testing positive for carp as evidence to update the model, with each new observation improving its robustness and its characterisation of uncertainty. However, the results can be unwieldy and difficult to present and interpret as the model is constructed, so Dr Schultz used additional statistical techniques to refine the model at each stage without losing its strength. As a consequence, he also obtained measures of the credibility of the eDNA concentration calculated for each region of the waterway.

…

When comparing the eDNA concentration estimates from different regions of the waterway, it is difficult to determine which estimates are from well-sampled regions versus those from poorly-sampled regions. The Bayesian inference model places equal weight on all evidence, regardless of the sampling frequency or intensity in that region. Evidence based on a small number of water samples or fewer sampling events can be regarded as weaker than evidence based on a larger number of samples or a greater number of sampling events.

Dr Schultz used sensitivity analysis to identify which regions of the waterway had been sufficiently sampled. He introduced evidentiary criteria to exclude the weakest evidence from the iterative Bayesian updating procedure. He then observed the stability of the concentration estimates as he increased the stringency of those criteria.

By testing whether a derived eDNA concentration estimate is sensitive to dropping the weakest evidence from the Bayesian inference model, Dr Schultz could identify which results were not stable. He suggests that comparing the relative stability of estimates should allow researchers to determine the reliability of their results, because instability indicates that the frequency or intensity of sampling in a given area is insufficient.

Estimates of eDNA concentrations obtained using Bayesian inference are also influenced by the choice of a prior distribution, which is an initial characterization of uncertainty in eDNA concentration needed to begin the updating procedure. In the early iterations of updating, estimates can be heavily influenced by the prior, but this influence fades as the number of iterations increases.

Dr Schultz based the parameters of his prior distribution on a model of eDNA survey sensitivity, which suggested lower and upper bounds for the eDNA concentration. The influence of this prior distribution can be seen in the first few iterations of updating, suggesting a high degree of uncertainty in the concentration estimates. However, this uncertainty diminishes over the first few iterations of updating. While the prior has little influence on the final estimate of the concentration, it shows that concentration estimates based on too few iterations of updating may be biased.

…

While Dr Schultz has effectively demonstrated that combining genetic analysis with powerful statistical tools can provide a robust technique to determine eDNA concentrations in samples, he notes that there are limitations to the method that require further research.

For example, the Bayesian inference model is insensitive to the sequence in which new evidence is added. In his study, the consistency of the series of observations resulted in less uncertainty than a highly variable series of observations would have done. To analyse temporal trends, which might indicate changes in the distribution or strength of an eDNA source over time, he suggests that researchers could distribute the evidence over a sequence of distinct periods and update starting with a new initial prior distribution in each period. However, this approach would require a large number of sampling events.

In the future, the modelling tools developed by Dr Schultz could be used to gain valuable insights into the movements and population sizes of other animals, greatly aiding efforts to control invasive species, and protect vulnerable wildlife.

…

This SciPod is a summary of the paper 'Inference of genetic marker concentrations from field surveys to detect environmental DNA using Bayesian updating', from *PLOS One*.
https://doi.org/10.1371/journal.pone.0190603

For further information, you can connect with Martin Schultz at Martin.T.Schultz@usace.army.mil