Efficient Estimation of Amphibian Clutch Size Using Image Analysis of Compressed Globular Egg Masses

Andrew D. Moraga[1] and Erik Pervin

Geomatics and Landscape Ecology Research Laboratory, Department of Biology, Carleton University,
1125 Colonel By Drive, Ottawa, Ontario, Canada K1S 5B6
[1]Corresponding author, e-mail: andrewd.moraga@gmail.com

Abstract.—Clutch size estimation for amphibians is a time intensive task where researchers often opt to hand count eggs one at a time, limiting the application of clutch size information in research. Contemporary methods seek to build a standard curve from a representative hand counted sample, then estimate the remaining samples based on measures such as displacement or dry weight. These methods, however, often result in the mortality of sampled egg masses while limiting the scalability and application of research connected to amphibian clutch size. Here we demonstrate the feasibility of an efficient and nondestructive survey method that integrates photographs taken in the field of compressed amphibian egg masses in an ovagram with open source computer software to automate counting in the lab. We compared manual counts to automated counts from compressed egg mass photographs of *Lithobates sylvaticus* (Wood Frog), *L. pipiens* (Northern Leopard Frog), and *Ambystoma maculatum* (Spotted Salamander), which were turned into black and white images with two ImageJ processes and then analyzed. We found that automated counts were much quicker than manual counts for all species and were more accurate for *L. sylvaticus* and *A. maculatum* than for *L. pipiens*. The main factor limiting accuracy in photograph counts was the resolution of individual eggs in an egg mass, suggesting future work should focus on greater attention to photographic conditions in the field.

Key Words.—anuran; fecundity; ImageJ; *Lithobates sylvaticus*; monitoring; reproduction; sampling techniques; Wood Frog

Introduction

Robust estimates of fecundity (i.e., clutch size) in many amphibian species are sparsely reported in the literature given the large time commitments necessary to obtain accurate measures. Current methods of estimating clutch size involve either some form of manual counts for individual eggs (Green 2015; Vershinin et al. 2015) or estimation through construction of a standard curve based on a sample of manually counted egg masses (Yerkmokhin et al. 2016). Manually counting eggs within a mass is time intensive due to the size of amphibian clutches (100s to 1000s of eggs per mass), thus highlighting a need for more efficient techniques to estimate clutch size.

Although contemporary methods used to estimate amphibian clutch size are more efficient than manually counting each egg with a globular mass, they are often destructive, either indirectly due to removal of the egg mass from its original habitat or through direct mortality from the counting technique. Direct sources of mortality can be attributed to methods of clutch size estimation that require desiccation to measure dry weights of masses (Yerkmokhin et al. 2016) and removal of the gelatinous matrix surrounding the eggs to incorporate either volumetric (Mitchell and Prague 2014) or manual counting. Direct mortality to the surveyed populations can limit the application of such methods, and the adoption of less destructive methods may promote more widespread counts and enable researchers to more feasibly scale up projects.

Karraker (2007) introduced a technique using an apparatus known as an ovagram allowing for non-destructive field estimates of clutch size for amphibians with globular egg masses. This technique involves compressing an amphibian egg mass between two transparent basins, one of which contains a grid to facilitate a systematic count of the number of eggs within each cell. Compressed egg masses are then returned, near to or exactly in their original position, which makes possible their continued development and survival. In a comparison of compressed and non-compressed egg masses, there was no difference found between the survivals of the two groups, suggesting that the ovagram method is a non-destructive technique to estimate amphibian clutch size.

The ovagram of Karakker (2007) is an effective tool to estimate amphibian clutch size in the field but it still necessitates time intensive manual counting of egg masses. Our objective was to demonstrate the feasibility of an integrated approach to amphibian clutch size estimation, which would decrease field survey time while increasing the accuracy and consistency of estimates across observers. We used the ovagram

to photograph compressed Wood Frog (*Lithobates sylvaticus*), Northern Leopard Frog (*Lithobates pipiens*), and Spotted Salamander (*Ambystoma maculatum*) egg masses, and ImageJ computer software adapted from cell biology, to estimate clutch size.

Materials and Methods

Our study region in Eastern Ontario, Canada, is characterized by mixed hard wood forests, seasonal and permanently flooded wetlands, and agricultural land. We surveyed 34 ponds (mean area 1,730 m^2, range 149–6,636 m^2) within the known distribution of *L. sylvaticus* and found their egg masses at 21 ponds. We found *L. pipiens* egg masses at seven ponds, and *A. maculatum* masses at six ponds. We used a double observer method for the egg mass surveys (as per Grant et al. 2005) from 13 April to 18 May 2016. We compared both manual counts and automated counts for both High Quality and Random Quality photographs of *L. sylvaticus* using linear regression in R (R Development Core Team 2017). We compared automated versus manual counts for *L. pipiens* and *A. maculatum* using a random set of photographs with linear regression.

We gently removed each egg mass from the water, placed it into a transparent flat-bottomed basin, and then gently compressed the mass between a second transparent basin to make the eggs individually distinguishable (as per Karraker 2007). The grid on the second basin was omitted because our method removes the need for manual counting in the field. We then took a photograph of the compressed egg mass with an Olympus Stylus TG-4 camera. We placed each mass back in the water next or near to where it was originally located. We uploaded the egg mass photographs to ImageJ, an open source image processing program (Schindelin et al. 2015). We manually counted the eggs in each egg mass using the ImageJ Multi-point feature, where a point was manually placed on each egg in the photograph; these were known as the manual counts. The software then tabulated the number of points that were present in the image.

We selected a subset of 45 High Quality photographs of compressed egg masses of *L. sylvaticus* through consideration of clarity and sharpness, minimal egg overlap, and consistent lighting conditions present within photographs. We uploaded and processed these photographs in ImageJ to count the number of eggs present within an egg mass based on a set of defined criteria, these were known as the automated counts. We selected a second subset of 59 Random Quality *L. sylvaticus* egg mass photographs from the total pool of photographs collected during the field survey. We used Microsoft Excel to select Random Quality photographs; we generated a random number beside each photograph ID using the RAND function, sorted the numbers from

low to high, and then selected the first 59 photographs. We then counted eggs within the Random Quality *L. sylvaticus* compressed egg mass photographs using the automated method.

For automated counts we used the Analyze Particles function to count the number of eggs present within a photograph. Analyze Particles requires a binarized (black and white) photograph to work. We achieved egg mass photograph binarization in two ways: using the Adjust Threshold process and the Find Maxima processes. For the Adjust Threshold process, we processed each egg mass photograph through several steps in the program before estimation via the automated method. First, we transformed an uploaded photograph into an 8-bit image (Image>8-bit). Once grayscale, we subtracted the background of the image to reduce noise (Process>Subtract Background): we selected light background and, depending on the photograph, we used a rolling ball radius between 5–10. Then we used the Threshold function (Image>Adjust>Threshold) to binarize the image. We used the default threshold settings in nearly all cases, with minor adjustments performed as needed. Lastly, we made the egg mass a region of interest (ROI) with the freehand polygon tool.

We reduced the remaining noise with the Remove Outlier function (Process>Noise>Remove Outliers): we used a threshold of 50 and Radius between 1 and 15 on Dark outliers. We applied the Fill Holes and Watershed function to the photographs to distinguish contiguous and overlapping eggs (Process>Binary>Fill Holes; Process>Binary>Watershed). Finally, we counted the eggs within the photograph using Analyze Particles (Analyze>Analyze Particles) with size variable, circularity between 0.3–1.00 and Show: Overlay. We adjusted the size until the overlay appeared to include all the recognizable eggs and we tabulated a final count with the Display Results function in analyze particles.

For the Find Maxima process, we needed fewer steps to binarize the photograph and estimate the number of eggs using Analyze Particles. We first applied the Find Maxima function (Process>Find Maxima): Noise tolerance was set to 50; Output type = Maxima Within Tolerance; Light background was checked. We then made the egg mass a ROI using the Oval selections tool. Finally, we applied Analyze Particles (Analyze>Analyze Particles): Size = 250–8000; Circularity = 0.20–1.00; Show = Overlay; checked Summarize. We applied the Watershed and Fill Holes functions if necessary. We adjusted the size until the overlay appeared to include all the recognizable eggs and we tabulated a final count with the Display Results function in analyze particles.

We compared automated count estimates and manual counts for the same High Quality and Random Quality photographs of egg masses of *L. sylvaticus* using linear regression in R. We considered the manual counts to reflect the actual clutch size as each individually

TABLE 1. Mean percentage difference with standard deviation between automated counts using the adjust threshold or find maxima binarization processes and the associated manual counts for each sample performed on photographs of compressed globular egg masses of Wood Frogs (*Lithobates sylvaticus*), Northern Leopard Frogs (*Lithobates pipiens*), and Spotted Salamanders (*Ambystoma maculatum*) using ImageJ. Abbreviations HQ = High Quality and RQ = Random Quality.

Species	Adjust Threshold (%)	Find Maxima (%)
Lithobates sylvaticus[HQ]	-2.61 ± 10.05	-5.07 ± 7.57
L. sylvaticus[RQ]	9.11 ± 15.70	-23.80 ± 14.77
L. pipiens	-2.71 ± 26.15	-58.98 ± 18.25
Ambystoma maculatum	20.73 ± 17.28	-3.45 ± 60.27

TABLE 2. Mean time (s) required with standard deviation for counting eggs of of Wood Frogs (*Lithobates sylvaticus*), Northern Leopard Frogs (*Lithobates pipiens*), and Spotted Salamanders (*Ambystoma maculatum*) within a given compressed globular egg mass photograph using an automated method with binarization through either the adjust threshold or find maxima processes and manual counts, only for *Lithobates sylvaticus*, performed using ImageJ. Abbreviations HQ = High Quality and RQ = Random Quality.

Species	Adjust Threshold (s)	Find Maxima (s)	Manual (s)
Lithobates sylvaticus[HQ]	127 ± 54	25 ± 10	354 ± 162
L. sylvaticus[RQ]	101 ± 40	19 ± 4	—
L. pipiens	97 ± 23	18 ± 4	—
Ambystoma maculatum	68 ± 20	17 ± 6	—

distinguishable egg was systematically counted by a person. We selected two random subsets of 15 egg mass photographs, as described above, for *L. pipiens* and *A. maculatum* as there was a much smaller pool of photographs available to choose from. We applied the automated method to estimate the number of eggs present within the *L. pipiens* and *A. maculatum* egg mass photographs as described above and compared them to the associated manual counts for each species using linear regression in R.

RESULTS

The time required to photograph a single compressed egg mass in the field from egg mass retrieval to replacement in the water was approximately 30 s for all species. The mean ± (SD) eggs per egg mass by hand count was: 501 eggs ± 196 eggs (range, 27–966 eggs; n = 45) for *L. sylvaticus*, 2,264 eggs ± 649 eggs (range, 1,104–3,304 eggs; n = 15) for *L. pipiens*, and 76 eggs ± 43 eggs (range, 17–180 eggs; n = 15) for *A. maculatum*. The number of eggs estimated from photographs binarized through the Adjust Threshold process for Random Quality photographs of *L. sylvaticus* was strongly related to the associated manual counts ($F_{1,57} = 210.3$, $P < 0.001$, $r^2 = 0.79$) while those egg mass photographs processed using the Find Maxima function were not ($F_{1,57} = 73.68$, $P < 0.001$, $r^2 = 0.56$; Table 1). The number of *L. sylvaticus* eggs estimated using the Adjust Threshold process for the High Quality photographs were strongly related to the associated manual counts ($F_{1,47} = 578.8$, $P < 0.001$, $r^2 = 0.92$) and even more strongly related when the Find Maxima process was applied ($F_{1,44} = 964.2$, $P < 0.001$, $r^2 = 0.96$; Table 1). The Adjust Threshold process required much more time than Find Maxima for both Random Quality and High Quality photographs (Table 2).

For *L. pipiens*, the automated method was not a good estimate of the number of eggs present within an egg mass when compared to the associated manual counts for photographs binarized using either the Adjust Threshold process ($F_{1,13} = 16.19$, $P = 0.001$, $r^2 = 0.55$) or the Find Maxima process ($P = 0.433$; Table 1). The automated method required more time to count *L. pipiens* eggs present within a photograph binarized using the Adjust Threshold process than for those binarized using the Find Maxima process (Table 2). For *A. maculatum*, the automated method was a good estimate of the number of eggs present within an egg mass based on the associated manual counts for photographs, for photographs binarized using either the Adjust Threshold process ($F_{1,13} = 70.35$, $P < 0.001$, $r^2 = 0.84$) or the Find Maxima process ($F_{1,13} = 34.08$, $P < 0.001$, $r^2 = 0.72$; Table 1). The automated method required more time to count the *A. maculatum* eggs present within a photograph binarized using the Adjust Threshold than for those binarized using the Find Maxima process (Table 2).

DISCUSSION

Our findings suggest that automated counts of *L. sylvaticus* and *A. maculatum* egg masses photographed in an ovagram are an effective non-destructive method for quickly and accurately estimating clutch size for these two species, but for *L. pipiens* automated counts do not provide acceptable estimates of clutch size. Photographs of compressed *L. pipiens* egg masses were challenging to process because the small distance between the eggs caused difficulties with resolving the boundaries between individual eggs and led to greater variation between the automated and manual counts. Photographs of *L. sylvaticus* and *A. maculatum* compressed egg masses were easier to process because the thicker jelly results in relatively larger distances between individual eggs, allowing for clear resolution when processed in the ImageJ software.

When examining the differences between photographs binarized using the Adjust Threshold and Find Maxima
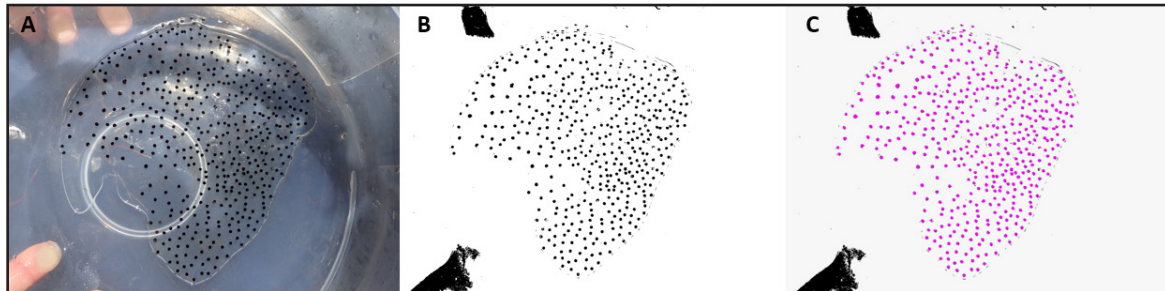
Figure 1. Photographs demonstrating the binarization process and automated counting for a high-quality photograph of an egg mass of a Wood Frog (*Lithobates sylvaticus*). The unprocessed egg mass photograph obtained in the field during an egg mass survey shows even light across the field of interest (A). The resulting binarized photograph demonstrates the benefit of high photographic quality with all eggs identified during the binarization process (B). The automated counts in the binarized photograph are shown in magenta and all eggs from the original photograph appear to have been counted (C). (Photographed by Erik Pervin).

processes, it becomes clear that there are circumstances when one method seems more appropriate than the other based on photograph quality. For *L. sylvaticus* egg masses, it appears that high quality photographs benefit from the greater accuracy of the Find Maxima process and the significantly lower amount of time required to estimate the clutch size of a given egg mass. Yet, this benefit does not appear to hold for Random Quality photographs binarized using the Find Maxima process. Despite the major time savings achieved with the Find Maxima process, it was noticeably less accurate than the Adjust Threshold process. In this study photograph quality was the driving factor in determining which process is more appropriate for estimating clutch size in *L. sylvaticus* egg masses.

Neither of the binarization processes is effective for automated counts of *L. pipiens*, a species with thinner jelly separating individual eggs. It is important, however, to highlight that most of the egg mass photographs for this species were obtained from a single site where all photographs suffered from poor local conditions and overhead cloud interference. For *A. maculatum*, egg mass photographs binarized with the Adjust Threshold process were more accurate than the Find Maxima process when estimating clutch size; however, there are major time savings from the Find Maxima process relative to Adjust Threshold. It is possible that more accurate estimates of clutch size could be obtained for *L. pipiens* and *A. maculatum* using the automated count method presented here if a larger sample size were used and photographic quality was improved relative to our study. Due to low occurrence of both species in the study region, proportionally fewer egg mass photographs were obtained compared to *L. sylvaticus*.

Regardless of the process selected to binarize photographs for the automated counts, the time savings are considerable. One *L. sylvaticus* egg mass photograph takes a mean of 354 s to count by hand, which allows for approximately 71 photographs to be counted over an 8-h work day, assuming a one-hour

break divided as desired. Compared to automated counts of high quality photographs, Adjust Threshold would require a mean of 127 s or 198 photographs per 8-h work day, and Find Maxima would require a mean of 25 s or 1,008 photographs per 8-h work day, both assuming a one-hour break divided as desired. The automated count estimates therefore provide significant time savings relative to manual counts. These tools also dramatically reduce the time in the field that is required for obtaining estimates of fecundity and would readily allow for increased research connected to clutch size in *L. sylvaticus* and *A. maculatum*.

While the automated method provided an effective estimate of clutch size for both *L. sylvaticus* and *A. maculatum*, the difference between the strength of relationship for Random Quality and High Quality photographs of *L. sylvaticus* compressed egg masses demonstrates that photograph quality is an important factor in the accuracy of automated counts. Ideal lighting conditions are the key determining factor of photograph quality. For the automated counting method presented in this paper, ideal conditions consist of high contrast between the eggs and their background and low glare from the surface of the upper transparent basin in the ovagram, which results in consistent lighting across the field of interest (Fig. 1). For ImageJ to properly binarize photographs, all the items of interest, individual eggs, in this case, must be of a similar value when converted to an 8-bit image. Failure to do so results in the loss of data, as binarizing eliminates all values below a certain threshold (Fig. 2). A bright spot on the photograph will wash out the data within that area, making it impossible to retain those points of interest when analyzing the photograph through Analyze Particles, and a dark spot will create a giant black region in the image, obscuring data or introducing large amounts of noise.

Most of the inaccuracies with the automated method arise from two sources: loss of data when binarizing and false positives from noise. Binarization necessarily eliminates data and the task of the image analyst is to
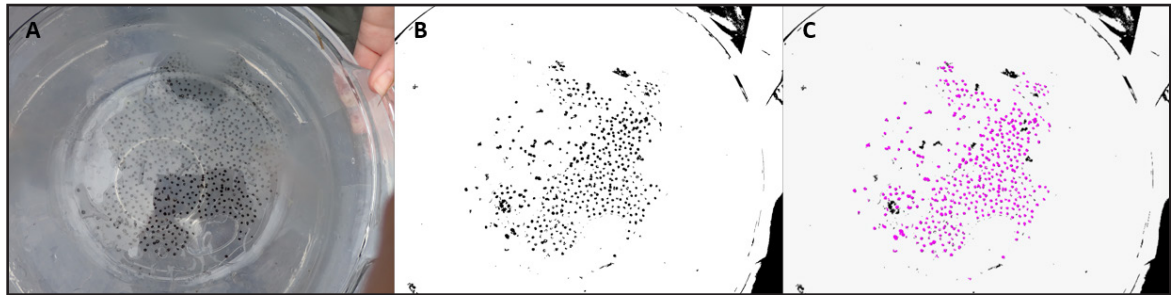
Figure 2. Photographs demonstrating the binarization process and automated counting for a low-quality photograph of an egg mass of a Wood Frog (*Lithobates sylvaticus*). The unprocessed egg mass photograph obtained in the field during an egg mass survey shows uneven lighting across the field of interest (A). The resulting binarized photograph demonstrates the inaccuracies in digitally separating eggs from the background in low quality photographs, with several missing (B). The automated counts in the binarized photograph are shown in magenta, with several missing (C). (Photographed by Erik Pervin).

minimize or eliminate the loss of data from regions of interest. This can only be done by starting with good quality photographs, suggesting that image enhancement is not often a practical surrogate for proper image acquisition. Although false positives are a function of photograph quality, they can also be produced through image enhancement, for example with the Watershed function. Watershed is meant to segment eggs that are overlapping or contiguous. Instead of properly separating two or more eggs, however, it can convert overlaps into dozens, if not hundreds, of false positives. This can be controlled to some extent by adjusting settings like circularity or size in Analyze Particles, as necessary, but this is time intensive and introduces potential for more human error. Avoiding the Watershed function is only useful for photographs that are already well segmented and therefore do not require it, and photographs that may benefit from Watershed tend to have noise which the Watershed function amplifies. It may be that the loss of data from binarization and the false positives from Watershed and other noise may somewhat cancel out, although beginning with good photograph quality goes a long way to eliminating both factors and greatly speeding up analysis at the same time. Ensuring that eggs are well segmented (not overlapping) is an important factor for image analysis. Therefore, it is important to control for non-ideal photography conditions as much as possible or practical. Future researchers employing this method should focus on improving photograph quality when applying this integrated method in the field.

Further improvements to photograph quality in this integrated method could be realized by improving photography standards. We recommend exploring the construction of an apparatus to maintain inter-photographic consistency of lighting, angle, and distance between camera and egg mass. For example, an arm of standard length that can be connected to the ovagram and has the camera affixed to it with a barrier around the camera to block obtrusive reflections from above. Secondly, a plain, opaque, contrasting background, such as one white flat-bottomed basin and one transparent flat-bottomed basin would be superior to the two transparent basins used in this study and better resolve the eggs within a given egg mass. Consistent photograph quality would allow for a single set of values to be used as the defined criteria and allow for full automation of the counting method, but so far this has only been achieved in a controlled laboratory setting with individual eggs not bound in a membrane (Bohenek and Resetarits 2017).

Although the current automated method is not fully automated in the fullest sense of the word, further improvements to photograph quality and consistency could lead to a single set of parameters applied across all photographs leading to batch upload counting. This would vastly reduce the time required to obtain clutch size estimates and, by extension, greatly reduce the cost of performing this kind of research. With the Find Maxima binarization process, we were almost able to achieve this, as it produced a drastic reduction in human input to analyze photographs when compared with the Adjust Threshold process. Here we have demonstrated the feasibility of a non-destructive method to estimate clutch size for amphibians through integration of the ovagram (Karraker 2007) with photography and open source computer software. This integrated technique allows for the sampling of large populations of globular egg mass laying amphibians across multiple survey locations and could vastly increase the amount of information available on reproductive ability in amphibians.

Literature Cited

Bohenek, J., and W.J. Resetarits. 2017. An optimized method quantify large number of amphibian eggs. Herpetology Notes 10:573–578.

Grant, E.H.C., R.E. Jung, J.D. Nichols, and J.E. Hines. 2005. Double-observer approach to estimating egg mass abundance of pool-breeding amphibians. Wetlands Ecology and Management 13:305–320.

Green, D.M. 2015. Implications of female body-size variation for the reproductive ecology of an anuran amphibian. Ethology Ecology & Evolution 27:173–184.

Karraker, N.E. 2007. A new method for estimating clutch sizes of ambystomatid salamanders and ranid frogs: introducing the Ovagram. Herpetological Review 38:46–48.

Mitchell, J.C., and C.A. Prague. 2014. Filling gaps in life-history data: clutch sizes for 21 species of North American anurans. Herpetological Conservation and Biology 9:495–501.

R Development Core Team. 2017. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Schindelin, J., C.T. Rueden, M.C. Hiner, and K.W. Eliceiri. 2015. The ImageJ ecosystem: an open platform for biomedical image analysis. Molecular Reproduction and Development 82:518–529.

Vershinin, V.L., S.D Vershinina, D.L. Berzin, D.V. Zmeeva, and A.V. Kinev. 2015. Long-term observation of amphibian populations inhabiting urban and forested areas in Yekaterinburg, Russia. Scientific Data 2: https://www.nature.com/articles/sdata201518.

Yerkmokhin, M.V., V.G. Tabachishin, and G.A. Ivanov. 2016. A method for estimating fecundity in the spadefood toad, *Pelobates fuscus*, through full and partial egg counting. Salamandra 52:337–341.
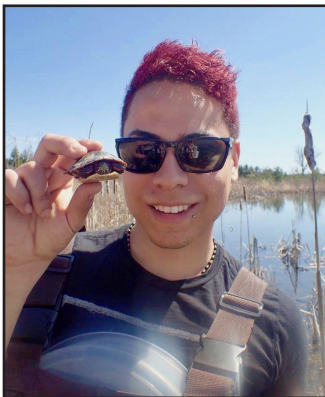
Andrew Moraga studied Environmental Science at Carleton University, Ontario, Canada, where he obtained a B.Sc. and went on to receive his M.Sc. in Biology also from Carleton University, where his research focused on amphibian conservation. He is now working towards a Ph.D. at the University of Waterloo, Ontario, Canada, examining novel techniques to rapidly restore mature hardwood forests on disturbed and degraded lands. He has a broad interest in using various conservation and restoration tools to contribute towards the ongoing effort of mediating biodiversity and habitat loss in a world with changing climate. (Photographed by Margaret Sawatzky)

Erik Pervin received his Bachelor of Humanities from Carleton University, Ontario, Canada. There he worked as a member of the Geomatics and Landscape Ecology Lab. His research interests range from the ecological implications of anthropogenic landscape changes to the philosophy of science. He is currently working outside of academia in the private sector. (Photographed by Andrew Moraga)