Herpetological Journal

Comparisons of image-matching software when identifying pool frog (*Pelophylax lessonae*) individuals from a reintroduced population

Josh Dawson[1], Connor T. Panter[1] & Inga Zeisset[1]

[1]Ecology, Conservation and Zoonosis Research and Enterprise Group, School of Pharmacy and Biomolecular Sciences, University of Brighton, Brighton, UK

Photographic identification of individual animals is a non-invasive and cost-effective method that can provide demographic information on wild populations. This study aims to compare two photo-matching algorithms (Wild-ID and I3S-Spot) using a reintroduced population of pool frogs (*Pelophylax lessonae*) in the UK as a case study. We compared the following parameters 1) sex and age, 2) image quality, 3) image collection size and 4) processing time to evaluate successful image match rates. There were no significant differences in successful match rates found between sex and age groups. Wild-ID was more sensitive to image quality than I3S-Spot. There was a significant negative relationship between image collection size and successful match rates for I3S-Spot, however, no such relationship for Wild-ID. The findings of our study can be used by conservation practitioners to reduce workload and improve accuracy during population monitoring activities.

Keywords: Amphibia, capture-recapture, I3S, photo identification, population monitoring, Wild-ID

Amphibians are experiencing global population declines due to threats such as climate change, habitat loss and infectious diseases (Antwis et al., 2014) and there is a need for adequate monitoring of wildlife populations to guide conservation activities. Methods including capture-mark-recapture (CMR) and capture-recapture (CR), are commonly used to assess population parameters in naturally occurring and reintroduced populations (Lagrange et al., 2014). Individuals are either identified by unique skin markings or artificial markers such as pit-tags, however, these methods can be time-consuming and cost ineffective (Bendick et al., 2013; Sannolo et al., 2016, Guimaraes et al., 2014; Griffiths et al., 2015; Chevalier et al., 2017).

The pool frog (*Pelophylax lessonae* - previously *Rana lessonae*) Camerano, 1882 is distributed across Europe but went extinct in the United Kingdom during the 1990s. Since then a reintroduction program has established a viable population at sites in England using individuals from the northern-clade in Sweden (Zeisset & Hoogesteger, 2018). Globally, the species is listed as 'Least Concern' on the IUCN Red List of Threatened Species but has a declining population trend and a need for conservation intervention has been recognised (Kuzmin et al., 2009). Pool frogs have distinctive and variable spotted skin patterns with a pale dorsal stripe, and exhibit some sexual dimorphism whereby adult males have paler dorsal base-colours than females (Hoogesteger et al., 2013). The presence of distinguishable morphological features and a need for adequate and minimally invasive population monitoring, mean that pool frogs are a suitable species for non-invasive photographic monitoring techniques. Photographic records of individuals from the reintroduced UK population are currently captured throughout the year, from which population estimates are currently calculated based on CR techniques (John Baker, pers. comm). However, this process is very time consuming for conservation practitioners.

There are now several automated software algorithms available that can aid this process and have been employed across a range of taxa, including amphibians (Elgue et al., 2014; Sannolo et al., 2016; Matthé et al., 2017; Speybroeck & Steenhoudt, 2017; Patel & Das, 2020), reptiles (Treilibs et al., 2016), large terrestrial mammals such as the Thornicroft's giraffe (*Giraffa cameleopardalis* thornicroft) (Halloran et al., 2014) and several species of elasmobranch (Gonzalez-Ramos et al., 2017; Navarro et al., 2018). However, such algorithms require varying amounts of input by the user and image quality (Yoshizaki et al., 2009; Halloran et al., 2014; Urian et al., 2015; Treilibs et al., 2016; Gonzalez-Ramos et al., 2017; Matthé et al., 2017; Gatto et al., 2018), data-set parameters (Matthé et al., 2017), morphological features of the study species (Yoshizaki et al., 2009; Urian et al., 2015; Matthé et al., 2017) and performance rates differ between software types.

This study aims to compare two commonly used photo-matching algorithms on a reintroduced population of pool frogs. We quantify differences between the algorithm types by 1) sex and age, 2) image quality, 3) image collection size and 4) processing time with respect

to successful image match rates. Temporal efficiency is important for population monitoring activities (Iijima et al., 2013) and we aim to make inferences on which algorithm practitioners should use to achieve maximum accuracy with the least amount of user effort.

To achieve this, a total of 1255 pool frog images taken between May and August from 2010 to 2017, were visually examined and 37 % (465) were deemed suitable as they had previously been identified to the individual-level by a species expert and for the purposes of this study, are accepted as being known matches. These images were also visually examined prior to analysis ensuring all were of appropriate quality and key morphological features were identifiable (examples of good quality and poor quality photographs available, see Supplementary Material Fig. S1). Reasons for poor quality images included poor lighting, light reflection from the skin of the individual and obstruction due to submersion in water. Image matches were confirmed and collated by year, sex (male/female) and age (adult/juvenile), thus forming 'image collections'. To identify individuals, we used the characteristic spotted dorsal pattern which is common amongst the species and is unique at the individual-level (Hoogesteger et al., 2013). To control for potential operator error, we repeated the image matching process three times for all image collections and each software. We recorded the number of successful first-time matches along with the time taken to process each collection. Time-per-image was calculated by dividing collection processing time by image collection size. We defined 'image collection size' as the number of images in each collection. We used the quantitative measure dots-per-inch (DPI) to assess image quality (Zhang & Gourley, 2008) and basic summary statistics (mean, standard deviation and standard error) were calculated for each image collection. Successful match rates between the two algorithms were compared based only on images that presented with the highest match-likelihood score, as determined by each algorithm. We used a similar method to Sannolo et al. (2016) and calculated successful match rates as a percentage of known matches in each collection.

We compared the algorithms Wild-ID (Bolger et al., 2012; http://www.teamnetwork.org:8080/Wild.ID) and I3S-Spot (Hartog & Reijns, 2014; http://www.reijns.com/i3s) as they have previously been used in similar studies (see Bendick et al., 2013; Sannolo et al., 2016). Technical differences between algorithms are available in the Supplementary Material and Figure S2, and a step-by-step overview of our methods are available in Figure S3. Images were confirmed as matches if the first image presented by the algorithm was an identifiable match to the image it was being compared to. Overall match rates and time-per-image means were compared between algorithms using a paired t-test. Average match rate variance between sexes and ages was tested using a One-Way ANOVA. This was also analysed by algorithm, however, for Wild-ID we used a Kruskal-Wallis test due to non-parametric data. Simple linear regression was used to test for significant linear relationships between image quality (DPI) and match rates (overall vs. Wild-ID vs. I3S-Spot), image collection size and match rates (overall vs.

Wild-ID vs. I3S-Spot), and time-per-image and match rates (overall). All statistical tests were performed using SOFA Statistics version 1.4.6 using an α value of 5 %.

Overall, the mean successful match rate was significantly greater for I3S-Spot (54.1 % ± 0.1 (SE)) than Wild-ID (40.8 % ± 0.2) ($t_{1,50}$ = 4.528, P <0.001). The average processing time per image (seconds) was also significantly greater for I3S-Spot (35 ± 3.7) than for Wild-ID (6 ±1.5) ($t_{1,50}$ = 65.956, P <0.001).

There were no significant differences in successful match rates between images of adult females (49.7 % ± 0.2 (SD)), adult males (48.4 % ± 0.2) and juveniles (40.2 % ± 0.2) ($F_{1,2}$ = 2.327, P > 0.05). When filtered by software type, there were no significant differences between median successful match rates for images of adult males (43.8 %; range 16.7-55.6 %) than adult females (42.9 %; 25-77.8 %) and juveniles (33.3 %; 12.5-54.6 %) using Wild-ID ($H_{1,2}$ = 3.116, P > 0.05) or I3S-Spot (match rates for images of adult males 57.7 % ± 0.1 compared with adult females 53.7 % ± 0.2 and juveniles 46.7 % ± 0.1) ($F_{1,2}$ = 2.070, P > 0.05). An overview of pool frog demographics in each dataset is provided in Supplementary Table S1.

When filtered by algorithm, there was a siginificant positive relationship between mean DPI and successful match rate for Wild-ID ($R^2_{1,49}$ = 0.209, P <0.001; Fig. 1a). There was no significant relationship between mean DPI and successful match rate for I3S-Spot ($R^2_{1,49}$ = 0.018, P >0.05; Fig. 1b).

There was also no significant relationship between image collection size (number of images) and successful match rate for Wild-ID ($R^2_{1,49}$ = 0.007, P >0.05; Fig. 1c). However, there was a significant negative relationship between image collection size and successful match rate for I3S-Spot ($R^2_{1,49}$ = 0.065, P <0.05; Fig. 1d) with higher match rates for smaller collections.

When combining data from both algorithms, there was a significant positive relationship between processing time-per-image and successful match rate ($R^2_{1,100}$ = 0.083, P <0.01). Furthermore, there was no significant relationship between time-per-image and successful match rate for Wild-ID ($R^2_{1,49}$ = 0.019, P >0.05; Fig. 1e), however, there was a significant negative relationship for I3S-Spot ($R^2_{1,49}$ = 0.173, P <0.01; Fig. 1f).

This study is one of the first comparing between two commonly used photo-matching algorithms to monitor reintroduced pool frog populations. Our results are comparable to similar studies with sample sizes ranging from 92 to 852 (Elgue et al., 2014; Sannolo et al., 2016; Treilibs et al., 2016; Gonzalez-Ramos,et al., 2017; Navarro et al., 2018; Patel & Das, 2020). Unlike other studies, we applied minimal operator effort (considering only the first potential matches presented as correct or not and only selecting the minimum requirement of 12 feature points when using I3S-Spot) to achieve our findings whereas others used greater levels of user effort (Sannolo et al., 2016; Treilibs et al., 2016; Matthé et al., 2017). Whilst not compared directly in this work, it is recognised that automated image processing can achieve successful match rates at least equal to manual observation, with potential for an increase in accuracy in several cases (Bendick et al., 2013; Treilibs et al., 2016; Matthé et al., 2017; Pawley et al., 2018). When comparing processing
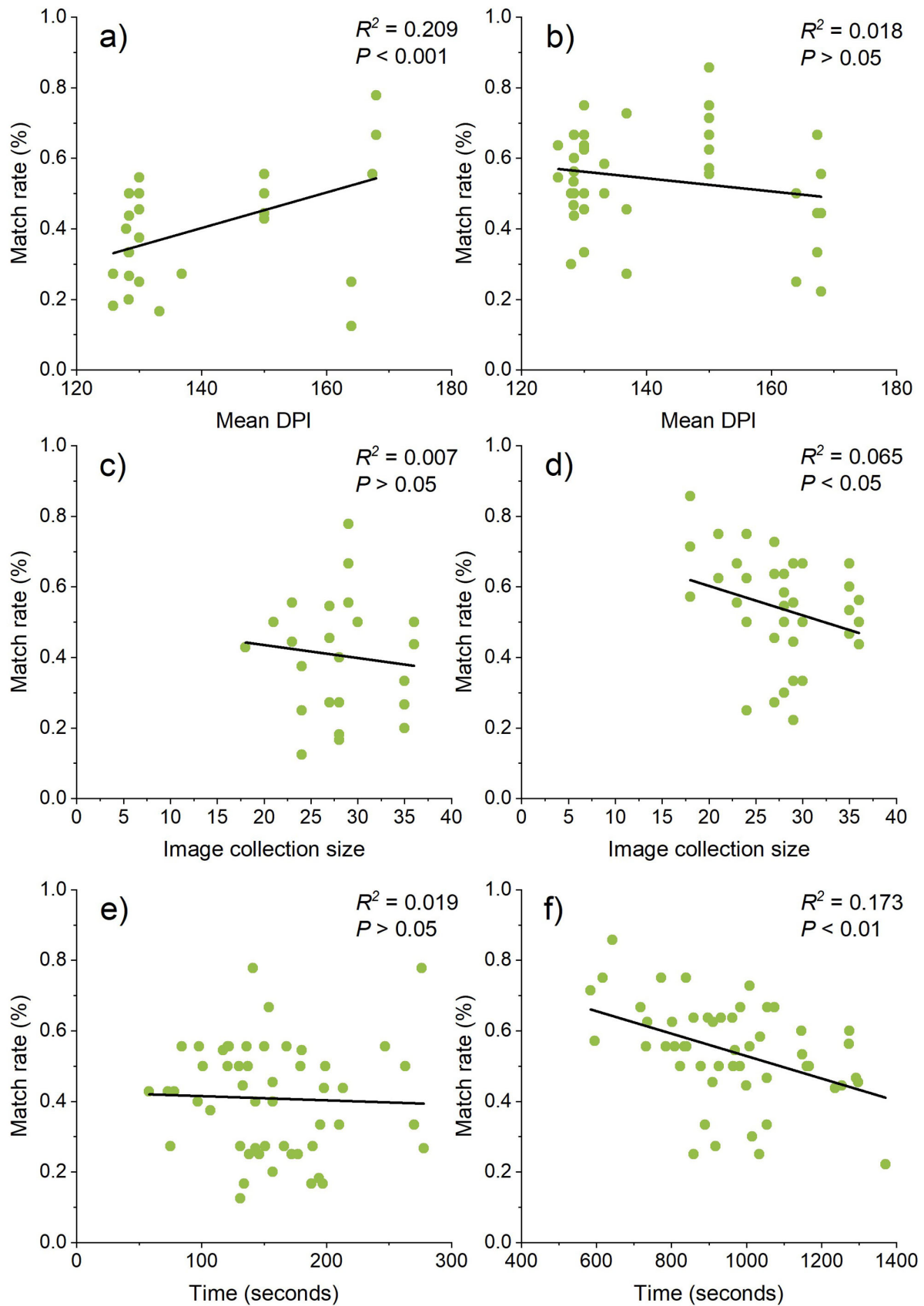
Table 1. Results from linear regression analyses a) Wild-ID match rate and mean image dots-per-inch (DPI), b) I3S-Spot match rate and mean image DPI, c) Wild-ID match rate and image collection size, d) I3S-Spot match rate and image collection size, e) Wild-ID match rate and time taken (seconds) and f) I3S-Spot match rate and time taken (seconds).

times, Wild-ID significantly outperformed I3S-Spot, with the latter taking more than six times longer on average to process images. However, the images used in this case were highly standardised due to a retrofit application of an existing image collection. It has been noted that feature-based algorithms, such as Wild-ID, benefit from image and pattern stability (Matthé et al., 2017) and it is likely that without such uniformity, the software would have performed less successfully. Despite this, we argue that in regards to time effciency both algorithms outperform manual efforts. This is similar to the findings of Halloran et al. (2014), who found that photo-id software could reduce processing time by up to 78 % compared with manual undertakings. We recommend users consider finding a balance between image processing time and successful match rates when using automated algorithms for accurate population assessments. Treilibs et al. (2016) recorded a 13 % increase in successful match rates when investing more time in image processing. We also found a positive relationship between image processing time and successful match rates for both algorithms.

Existing studies have revealed several influencing factors to be considered when using automated algorithms for photo-identification. Such aspects include changing features of indiviudals over time (Yoshizaki et al., 2009; Urian et al., 2015; Matthé et al., 2017), dataset size (Matthé et al., 2017) and image quality (Yoshizaki et al., 2009; Halloran et al., 2014; Urian et al., 2015; Treilibs et al., 2016; Gonzalez-Ramos et al., 2017; Matthé et al., 2017). Sex and age group did not significantly influence successful match rates. Other similar studies have found differences between the sexes, for example in Italian crested newts (*Triturus carnifex*) (Sannolo et al., 2016), however, there was no such difference in our study. This may be due to more extreme sexual dimorphism in Italian crested newts compared to pool frogs. There was a significant negative relationship between image collection size and successful match rates for I3S-Spot, which are similar to findings by Matthé et al. (2017). Our findings are also similar to other studies (Yoshizaki et al., 2009; Elgue et al., 2014; Halloran et al., 2014; Urian et al., 2015; Treilibs et al., 2016; Matthé et al., 2017; Gatto et al., 2018), however, these have been limited by qualitative approaches to assess image quality such as visual examination and manual scoring. It may be more appropiate to measure image quality via the user's ability to identify key morphological features, albeit this could be regarded as subjective and is dependent on user skill and effort. We used a quantitative approach in the form of DPI. Whilst DPI is a valid measurement of an image's quality, it does not classify an image based on the visibility of key features.

Our results have compared two existing automated photo-identification algorithms and highlighted differences between them. However, such algorithms are an asset when monitoring reintroduced populations allowing for the reduction in workload for conservation practitioners, but limited by differences between algorithms and identifying features between species. It is important that practitioners consider 1) an appropriate algorithm to use, 2) their project design and 3) balancing image processing time with successful match rates.

This will improve accuracy and efficiency of in-situ conservation and offer automated tools as a viable, if not preferable, alternative to other more-invasive and time-consuming monitoring techniques.

ACKNOWLEDGEMENTS

REFERENCES

Antwis, R. E., Purcell, R., Walker, S. L., Fidgett, A. L. & Preziosi, R. F. (2014). Effects of visible implanted elastomer marking on physiological traits of frogs. *Conservation Physiology* 2, DOI: 10.1093/conphys/cou042

Bendick, N. F., Morrison, T. A., Gluesenkamp, A. G., Sanders, M. S. & O'Donnell, L. J. (2013). Computer-assisted photo identification outperforms visible implant elastomers in an endangered salamander *Eurycea tonkawae*. *PLoS One* 8, e59494.

Bolger, D. T., Morrison, T. A., Vance, B., Lee, D. & Farid, H. (2012). A computer-assisted system for photographic mark-recapture analysis. *Methods in Ecology and Evolution* 3, 813-822.

Chevalier, H. L., Calvez, O., Martinez-Silvestre, A., Picard, D., Guerin, S. & Isselin-Nondedeu, F. (2017). Marking techniques in the Marbled newt (*Triturus marmoratus*): PIT-Tag and tracking device implant protocols. *Acta Herpetologica* 12, 79-88.

Elgue, E., Pereira, G., Achaval-Coppes, F. & Maneyro, R. (2014). Validity of photo-identification techniques to analyze natural markings in *Melanophryniscus montevidensis*. *Phyllomedusa* 13, 59-66.

Gatto, C. R., Rotger, A., Robinson, N. J. & Tomillo, P. S. (2018). A novel method for photo-identification of sea turtles using scale patterns on the front flippers. *Journal of Experimental Marine Biology and Ecology* 506, 18-24.

Gonzales-Ramos, M. S., Santos-Moreno, A., Rosas-Alquicira, E. F. & Fuentes-Mascorro, G. (2017). Validation of photo-identification as a mark-recapture method in the spotted eagle ray *Aetobatus narinari*. *Journal of Fish Biology* 90, 1021-1030.

Griffiths, R. A., Foster, J., Wilkinson, J. W. & Sewell, D. (2015). Science, statistics and surveys: a herpetological perspective. *Journal of Applied Ecology* 52, 1413-1417.

Guimarães, M., Corrêa, D. T., Filho, S. S., Oliveira, T. A. L., Doherty, P. F. & Sawaya, R. J. (2014). One step forward: contrasting the effects of toe clipping and PIT tagging on frog survival and recapture probability. *Ecology and Evolution* 4, 1480-1490.

Halloran, K. M., Murdoch, J. D. & Becker, M. S. (2014). Applying computer-aided photo-identification to messy datasets: a case study of the Thornicroft's giraffe (*Giraffa cameleopardalis* thornicroft). *African Journal of Ecology* 53, 147-155.

Hartog, J. D. & Reijns, R., (2014). I3S Spot Manual, June

2014, version 4.0.2. Available at: http://www.reijns.com/i3s/download/I3S_download.html Accessed 5 November 2018.

Hoogesteger, T., Rahkonenm J. & Karhilahti, A. (2013). Pool frog (*Pelophylax lessonae*) Camerano 1882 (Anura, Ranidae), an addition to the Finnish amphibian fauna. *Memoranda Societatis pro Fauna et Flora Fennica* 89, 25-31.

Iijima, H., Nagaike, T. & Honda, T. (2013). Estimation of deer population dynamics using a Bayesian state-space model with multiple abundance indices. *The Journal of Wildlife Management* 77, 1038-1047.

Kuzmin, S., Beebee, T., Andreone, F., Nyström, P., Anthony, B.P., Schmidt, B., Ogrodowczyk, A., Ogielska, M., Cogalniceanu, D., Kovács, T., Kiss, I., Puky, M. & Vörös, J. (2009). *Pelophylax lessonae* (errata version published in 2016). The IUCN Red List of Threatened Species 2009: e.T58643A86643256. https://dx.doi.org/10.2305/IUCN.UK.2009.RLTS.T58643A11818386.en. Downloaded on 04 May 2020.

Lagrange, P., Pradel, R., Belisle, M. & Gimenez, O. (2014). Estimating dispersal among numerous sites using capture-recapture data. *Ecology* 95, 2316-2323.

Matthé, M., Sannolo, M., Winiarksi, K., Spitzen – van der Sluijs, A., Goedbloed, D., Steinfartz, S. & Stachow, U. (2017) Comparison of photo-matching algorithms commonly used for photographic capture-recapture studies. *Ecology and Evolution* 7, 5861-5872.

Navarro, J., Perezgrueso, A., Barria, C. & Coll, M. (2018). Photo-identification as a tool to study small-spotted catshark *Scyliorhinus canicula*. *Journal of Biology* 92, 1657-1662.

Patel, N. G. & Das, A. (2020). Shot the spots: a reliable field method for individual identification of *Amolops formosus* (Anura, Ranidae). *Herpetozoa* 33, 7-15.

Pawley, M. D. M., Hupman, K. E., Stockin, K. A. & Gilman, A. (2018). Examining the viability of dorsal fin pigmentation for individual identification of poorly-marked delphinids. *Scientific Reports* 8, 12593.

Sannolo, M. Gatti, F., Mangiacotti, M., Scali, S. & Sacchi, R. (2016). Photo-identification in amphibian studies: a test of I3S pattern. *Acta Herpetologica* 11, 63-68.

Speybroeck, J. & Steenhoudt, K. (2017). A pattern-based tool for long-term, large-sample capture-mark-recapture studies of fire salamanders *Salamandra* species. *Acta Herpetologica* 12, 55-63.

Treilibs, C. E., Pavey, C. R., Hutchinson, M. N. & Bull, C. M. (2016). Photographic identification of individuals of a free-ranging, small terrestrial vertebrate. *Ecology and Evolution* 6, 800-809.

Urian, K., Gorgone, A., Read, A., Balmer, B., Wells, R. S., Berggren, P., Durban, J., Eguchi, T., Rayment, W. & Hammond, P. S. (2014). Recommendations for photo-identification methods used in capture-recapture models with cetaceans. *Marine Mammal Science* 31, 298-321.

Yoshizaki, J., Pollock, K. H., Brownie, C. & Webster, R. A. (2009). Modelling misidentification errors in capture-recapture studies using photographic identification of evolving marks. *Ecology* 90, 3-9.

Zeisset, I. & Hoogesteger, T. (2018). A reassessment of the biogeographic range of northern clade pool frogs (*Pelophylax lessonae*). *Herpetological Journal* 28, 63-72.

Zhang, A. B. & Gourley, D. (2008). 5 – Digitising material. In: Creating Digital Collections. Chandos Publishing, 55-72.

Please note that the Supplementary Materials are available via the Herpetological Journal website: https://thebhs.org/publications/the-herpetological-journal/volume-31-number1-january-2021