Madon et al 2011_MEE

1. Mark–recapture studies are often used to estimate population size based on a single source of individual identification data such as natural markings or artificial tags. However, with the development of molecular ecology, multiple sources of identification can be obtained for some species and combining them to obtain population size estimates would certainly provide better information about abundance than each survey can provide alone.

2. We propose an extension of the Jolly–Seber model to infer abundance by combining two sources of capture–recapture data. The need to merge both sources of data was motivated by studies of humpback whales in which both photo-identification and DNA from skin biopsy samples are often collected. As whales are not necessarily available by both sampling methods on any given occasion, they can appear twice in the combined data set if no combined sampling ever occurred during the survey, i.e. being photographed and genotyped on the same occasion. Our model thus combines the two sources of information by estimating the possible overlap. Monte Carlo simulations are used to assess the properties of the present estimator that is then used to estimate the size of the humpback whale population in New Caledonia. The new openpopulation estimator is also compared with classic closed-population estimators incorporating either temporal and∕ or individual heterogeneity in the capture probability: the purpose was to evaluate which approach (closed or open population) was the least biased for an open population with individual heterogeneous capture probabilities.

3. When all assumptions are met, the estimator is unbiased as long as the probability of being double-tagged (e.g. photographed and biopsied on the same occasion) on every occasion is above 0Æ2.

4. The humpback whale case study in New Caledonia shows that our two-source Jolly–Seber (TSJS) estimator could be more efficient in estimating population size than models based only on one type of data. For monitoring purposes, the proposed method provides an efficient alternative to the existing approaches and a productive direction for future work to deal with multiple sources of data to estimate abundance.

5. R-codes formatting the data and implementing the TSJS model are provided in Resource S5.