Family size dynamics in wintering geese

Online resource 01

Pratik R. Gupte, Kees Koffijberg, Gerard JDM Müskens, Martin Wikelski and Andrea Kölzsch

Handling data from www.geese.org

We obtained 18,035 observations from www.geese.org, the online portal to which citizen scientists may report resightings of neckbanded geese. From these, we:

1. Removed 3,360 observations where the neckband code of focal geese and that of their partner was reported to be the same.

2. Removed 1,675 observations where the birth year was unknown or less than two years prior to the date of being observed.

3. Dealt with reported family sizes in the following way:

   a) We found that observers had entered 73 unique values in the data field for number of juveniles (family size).

   b) These represented a range of certainty, from the authoritative (eg. `0,1,2, ... 9`, or `4, ongerind`) to the unsure (eg. `2 - 4`, or `minimaal 3, max.5`), to the completely uncertain (eg. `>0`, or `1 (minimum)`).

   c) We extracted the first two characters of each of these values (eg. family size reported `0` extracted as `0`; family size reported `10` extracted as `10`; family size reported `2 - 4` extracted as `2`; family size reported as `minimaal 3, max.5` extracted as `mi`).

   d) This resulted in a conservative (lower) estimate of family sizes in cases such as `2 - 4` (estimated as 2), and a loss of data in cases such as `minimaal 3, max.5` (considered as `missing data`). The latter case could have been rescued and assigned a value of 3, but no general rule could be found that suited all 73 cases.

   e) This process resulted in the removal of 405 observations, or 3.1% of the 13,000 observations remaining following steps 1 and 2.

4. This entire filtering process discarded 5,440 observations (30%) and retained 12,595 observations, which were then filtered as described in the methods for young birds and single birds without juveniles.