Methods: Data

Flock counts

We gained access to four datasets relating to the study species from our region of interest. The first was a set of counts of flock sizes, of the age ratio (percentage of first winter birds), of the numbers of families of each flock, and other associated information (time, habitat type, observer identity), made by volunteers ($n = 75$) across western and central Europe ($n = 8764$). These data only counted successful families, ie, the minimum family size was one (juvenile), and lacked information on the social (paired or not) and breeding status (unsuccessful or immature) of the remainder of the geese in the flock.

Flocks are often larger than can be sampled accurately in the field, leading to mistakes. In 6% of cases, the flock size was missing, and this was reconstructed as either the related (and in 31% of cases, identical) number of geese sampled, or, if that value was also missing ($n = 28$), as the sum of the number of adults and juveniles. In cases where the percentage of first winter birds was missing ($n = 50$), it was calculated from the number of juveniles and the flock size.

We filtered these data spatially, including only records from the Netherlands ($\sim$ 52%), and from the German state of North-Rhine Westphalia ($\sim$ 40%). We then filtered them temporally; though the data contained observations since 1957, the number of records before the breeding year 2000 was low ($\sim$ 7%), and we retained only records made between autumn 2000 and spring 2017 ($n = 7,416$). In order to facilitate downstream analyses, we checked whether each record had a complete date (year, month, day), and in cases where the day was missing ($n = 32$), we assigned it to the 15th. In all cases hereon, 'year' refers to the goose breeding year, which begins in June.

Each record was assigned one of the zones, "East Frisia", "IJsselmeer", "Rhinelands", and "Southwest", based on the province in the Netherlands, or the district in Germany, in which it was made. Zones cover areas in or around the names given. The number of families in each flock was counted, and a subset of the data ($n = 2014$) in which the number of families present in flocks was recorded was set aside for analyses relating to family size.

Nearly all (> 99%) records had associated site names, with flocks recorded at 1567 unique sites. This data was geocoded using the Google Maps Geocoding API accessed by the function geocode from the package ggmap (Kahle and Wickham 2013) using one of 123 unique look-up names associated with the site names. Exclusion of sites without geocoded coordinates, or with coordinates outside the box bounded by 0°E and 10°E, and 50°N and 54°N, left 7141 records of flocks, in 1884 of which family sizes and frequencies were also known. Family level data was then extracted from each flock to get variables for each family (n families = 51,037).

Observations of marked geese

We obtained data on sightings and positions of individual geese marked with numbered neckbands, and reported using tools available on geese.org, an initiative to track marked geese via direct observations [cite the website? pers. comm?]. We removed the following types of records: those in which a goose was seen with neither juveniles nor a social partner, those in which two birds of a pair had been reported separately when seen together, those in which the bird was less than two years old at the time of observation, and those which lay outside the bounding box and temporal range described above. 10,635 records remained, reported from 8,416 unique coordinate pairs.

These data differed from the flock level counts in three important respects. First, the records lacked information on the size of the flock in which each goose was seen, and on other flock attributes, such as habitat type. Second, the data contained records of pairs . Finally, the positions for each observation were nearly unique since observers recorded the sighting online using a drag-and-drop locator, as part of the functionality of `geese.org`.

Flight activity

To determine the approximate dates that bookend the stay of geese in their wintering grounds, we accessed goose flight activity data from Trektellen [cite website? pers. comm?] sites across the Netherlands. These data were filtered to exclude sites that lay close to goose night roosts so as to avoid noise in the data from daily, rather than migratory, movements. Data were further filtered to exclude flights that did not match the direction of migration appropriate to the season. From these data, we calculated the first date in each autumn (ending December) and the last date in each spring (beginning January) on which goose flight activity, in the form of number of geese flying per hour of observation time, was at or above the 90th percentile of within- breeding year season specific activity. We took these dates to represent the beginning of goose arrival from the autumn migration, and the end of goose departure on the spring migration, respectively. We added these dates to the flock, family, and individual level data, matching them by breeding year, and then calculated the number of days between each observation and the two extremes of goose migration.

Summer breeding success

The number of juveniles in families observed on the wintering grounds is a function of the breeding success of geese in the Arctic. This in turn is thought to be linked to the abundance of Arctic rodents, primarily lemmings (*Lemmus sp.* and *Dicrostonyx sp.*). This shows a cyclical pattern with a 3 – 4 year period, with a 'lemming peak' year followed immediately by a 'crash' year, with abun-

dance rising until the next peak. The factors underlying that are thought to be largely intrinsic, but it may also be related to the form of precipitation and its effect on the availability of vegetation to lemmings (Hansen et al. 2013).

Summers and Underhill (1987) hypothesised that goose breeding success is high in 'lemming peak' years, since Arctic predators switch from their preferred lemming prey to goose eggs and young when lemmings are scarce. Dhondt (1987) amended this alternate prey hypothesis (AHP) to reflect that predator populations are linked to, but also lag behind, lemming abundance. Goose breeding success is thus predicted to be lowest in the years immediately following lemming peaks, when a dearth of lemmings and a surfeit of predators combine to produce predation pressure on goose young that's higher than the mean. This logic has been found to explain winter estimates of breeding success of both geese and waders wintering in Scandinavia (Blomqvist et al. 2002), and of Dark-bellied Brent geese (*Branta b. bernicla*) wintering along the North Sea coast (Nolet et al. 2013).

Within this context, we sought to calculate an index for the summer predation pressure on our population, following a method presented earlier in Blomqvist et al. (2002), Nolet et al. (2013), and Koffijberg (2010). The breeding grounds of our population lie above the Arctic Circle (66.5°N), and between the Kanin Peninsula (45°E) and the River Yenisei (85°E) (Madsen and Cracknell 1999). We could not asses rodent abundance for this area from the literature, which focuses largely on lemming abundance from the Taimyr Peninsula (98°E) (Kokorev and Kuksov 2002, used in Blomqvist et al. (2002), and in Nolet et al. (2013)). While it is suggested that lemming cycles in some regions of Scandinavia (Angerbjörn et al. 2001), and in the high Arctic of Svalbard (Hansen et al. 2013) may be synchronised by climatic fluctuations, we did not expect lemming cycles to be synchronised with those of Taimyr, and preferred to use our own index.

We obtained rodent abundance indices for the relevant region from the website *Arctic Birds* `www.arcticbirds.net` [cite website?], an initiative of the International Breeding Conditions Survey on Arctic Birds. Nolet et al. (2013) previously used

the same approach to to fill in gaps in the dataset they used. Sites on *Arctic Birds* are marked on a map and binned into four classes: "Unclear", "Low or absent", "Average", and "High". The same sites are not present in each year. We graded these sites on a 0 - 3 scale, with 0 for "Unclear, and 3 for"High". The distinction between sites graded 0 and 1 was itself unclear. Both 0 and 1 were used to mark the island of Kolguyev, which is well known to have no lemming cycle [pers. obs.]. We took 0 to indicate a near or full absence of lemmings rather than an unsurveyed site, and also included an entry of 0 for Kolguyev in each year. Around 30% of the North Sea population of Whitefronts breeds on Kolguyev, where it forms the major prey base for the island's predators (Kondratyev and Zaynagutdinova 2008), making this an important addition to the data.

We averaged the lemming index across the sites in each year, and then for each year *t*, we calculated a predation index (P_t), which is highest in years immediately after lemming peaks, following Blomqvist et al. (2002) and Koffijberg (2010).

$$P_t = \frac{L_{t-1} - L_t + 3}{2} \tag{1}$$

Tracked families

The only families for which trends in size and position could be studied with absolute certainty were those which had been fitted with GPS receiving position logger/transmitters during the winters of 2013 ($n = 3$), 2014 ($n = 4$), and 2016 ($n = 6$). Position loggers in 2013 and 2014 were e-obs GmbH backpacks, and in 2016 were numbered neckband loggers supplied by Theo Gerrits (madebytheo). These loggers were set to record a baseline of one position every 30 minutes, though the actual fix frequency depended on the mode in which the remotely programmable device was then operating. Data from these loggers were uploaded remotely to the animal tracking database Movebank, from where they were retrieved prior to analysis.

Logger data were filtered to fit within the spatial extents of the study area, and data collected after March 31 each

spring were excluded from the analysis. A major component of these data were 'flight bursts', high fix frequency (0.5 or 1 Hz) records triggered by takeoff as measured by on-board accelerometers. These bursts were removed, and only data with the baseline sampling interval retained. Fixes where the logger position error was estimated to be above 20m were also removed. To fully account for irregularities in sampling interval introduced by the logger not functioning as ideal, the remaining data were averaged over every half hour so as to obtain a regular timeseries of data. This allowed for a meaningful matching of positions within families at the same timestamp.

The adult in the family with the greater number of logged positions was set to be the 'reference'. This aimed to obtain a longer sequence of reference-to-individual distances, which might also result in capturing more family dynamics. The distance between the reference and all other individuals in the family was calculated using the Vincenty ellipsoid method (Vincenty 1975) implemented by the geosphere package in R (Hijmans 2016). These distances were used to determine the number of family members within a 250m radius of the reference, and the family size per day was obtained as the maximum number of members within that radius during a day. Due to an accretion of errors at the level of the logger and in rounding and averaging the data, the sizes of some families as calculated above fluctuated drastically over time. We then considered the family size on each day to be maximum of the family sizes on all days between that one and the final day.

In one case, the entire family split with the male and female separating. We did not differentiate this case from other types of family size decrease, which in our data included juvenile independence, juvenile death, and logger malfunction.

———————————

References

Angerbjörn, A. et al. 2001. Geographical and temporal patterns of lemming population dynamics in fennoscan-

dia. – Ecography 24: 298–308.

Blomqvist, S. et al. 2002. Indirect effects of lemming cycles on sandpiper dynamics: 50 years of counts from southern sweden. – Oecologia 133: 146–158.

Dhondt, A. A. 1987. Cycles of lemmings and brent geese branta b. bernicla: A comment on the hypothesis of rose-laar and summers. – Bird Study 34: 151–154.

Hansen, B. B. et al. 2013. Climate events synchronize the dynamics of a resident vertebrate community in the high arctic. – Science 339: 313–315.

Hijmans, R. J. 2016. Geosphere: Spherical trigonometry.

Kahle, D. and Wickham, H. 2013. Ggmap: Spatial visualization with ggplot2. – The R Journal 5: 144–161.

Koffijberg, K. 2010. Breeding success amongst greater whitefronted geese in 2009/10–a progress report. – Goose Bull. Nov 2010: 32–34.

Kokorev, Y. and Kuksov, V. 2002. Population dynamics of lemmings, lemmus sibirica and dicrostonyx torquatus, and arctic fox alopex lagopus on the taimyr peninsula, siberia, 1960–2001. – Ornis Svecica 12: 139–143.

Kondratyev, A. and Zaynagutdinova, E. 2008. Greater white-fronted geese (anser albifrons) and bean geese (a. fabalis) on kolguev island–abundance, habitat distribution, and breeding biology. – Vogelwelt 129: 326–333.

Madsen, J. and Cracknell, G. 1999. Goose populations of the western palearctic. in press.

Nolet, B. A. et al. 2013. Faltering lemming cycles reduce productivity and population size of a migratory arctic goose species. – Journal of animal ecology 82: 804–813.

Summers, R. and Underhill, L. 1987. Factors related to breeding production of brent geese branta b. bernicla and waders (charadrii) on the taimyr peninsula. – Bird Study 34: 161–171.

Vincenty, T. 1975. Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations. – Survey review 23: 88–93.