Centennial relationships between ocean temperature and Atlantic puffin production reveal shifting decennial trends


It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1111/gcb.15665

Publisher: Wiley-Blackwell

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur
CentAUR
Central Archive at the University of Reading
Reading’s research outputs online
Centennial relationships between ocean temperature and Atlantic puffin production reveal shifting decennial trends

Erpur S. Hansen¹ | Hanno Sandvik²,³ | Kjell Einar Erikstad³,⁴ | Nigel G. Yoccoz⁴,⁵ | Tycho Anker-Nilssen² | Jürgen Bader⁶,⁷ | Sébastien Descamps⁸ | Kevin Hodges⁹ | Michel d. S. Mesquita⁷ | Tone K. Reiertsen⁴ | Øystein Varpe¹⁰,¹¹,¹²

¹South Iceland Nature Research Centre, Vestmannaeyjar, Iceland
²Norwegian Institute for Nature Research (NINA), Trondheim, Norway
³Centre for Biodiversity Dynamics (CBD), Norwegian University of Science and Technology (NTNU), Trondheim, Norway
⁴Norwegian Institute for Nature Research (NINA), Tromsø, Norway
⁵Department of Arctic and Marine Biology, Arctic University of Norway (UiT), Tromsø, Norway
⁶Max Planck Institute for Meteorology, Hamburg, Germany
⁷Bjerknes Centre for Climate Research, NORCE, Bergen, Norway
⁸Norwegian Polar Institute, Tromsø, Norway
⁹Department of Meteorology, University of Reading, Reading, UK
¹⁰Department of Biological Sciences, University of Bergen, Bergen, Norway
¹¹University Centre in Svalbard, Longyearbyen, Norway
¹²Norwegian Institute for Nature Research (NINA), Bergen, Norway

Abstract
The current warming of the oceans has been shown to have detrimental effects for a number of species. An understanding of the underlying mechanisms may be hampered by the non-linearity and non-stationarity of the relationships between temperature and demography, and by the insufficient length of available time series. Most demographic time series are too short to study the effects of climate on wildlife in the classical sense of meteorological patterns over at least 30 years. Here we present a harvest time series of Atlantic puffins (Fratercula arctica) that goes back as far as 1880. It originates in the world’s largest puffin colony, in southwest Iceland, which has recently experienced a strong decline. By estimating an annual chick production index for 128 years, we found prolonged periods of strong correlations between local sea surface temperature (SST) and chick production. The sign of decennial correlations switches three times during this period, where the phases of strong negative correlations between puffin productivity and SST correspond to the early 20th century Arctic warming period and to the most recent decades. Most of the variation (72%)
in chick production is explained by a model in which productivity peaks at an SST of 7.1°C, clearly rejecting the assumption of a linear relationship. There is also evidence supporting non-stationarity: The SST at which puffins production peaked has increased by 0.24°C during the 20th century, although the increase in average SST during the same period has been more than three times faster. The best supported models indicate that the population’s decline is at least partially caused by the increasing SST around Iceland.

**KEYWORDS**
chick production, *Fratercula arctica*, Iceland, long-term temperature trend, population decline, sea surface temperature

## 1 | INTRODUCTION

Changes in climatic conditions can have profound effects on the demography, phenology and population trajectories of marine top predators such as seabirds (Dias et al., 2019; Jenouvrier, 2013; Oro, 2014; Sydeman et al., 2015). In light of projected climatic change due to global warming (IPCC, 2013), this raises concerns for the viability of populations that respond negatively to increasing temperatures, either directly or mediated by effects at lower trophic levels (Jones et al., 2018; Sandvik et al., 2014; Trathan et al., 2020).

Thermal performance curves of species determine their responses to global warming, which can create a mismatch between the temperature optimum and the temperatures experienced (McKenzie et al., 2020). Such thermal performance curves are well documented in ectotherms, such as fish (e.g. van der Kooij et al., 2008). Effects at more basal levels of the trophic pyramid can, in turn, affect species at more apical trophic levels (Carroll et al., 2016).

According to the classical definition, climate is a statistical description of weather over a 30-year period (e.g. IPCC, 2013, p. 1450). Investigating the ecological effects of changes in climate therefore requires even longer datasets. However, the number of sufficiently long ecological time series is limited (but see e.g. Wilming et al., 2020). Such paucity of long-term time series hampers especially our understanding of climate effects on long-lived species, such as seabirds or other top-predators, for which most available time series only cover one or few generations. For instance, relationships between environmental parameters and wildlife populations are very likely neither linear nor stationary (e.g. Descamps et al., 2017; Mysterud et al., 2001), whereas non-linearity and non-stationarity are difficult to address with short time series. Consequently, relationships based on short-term data may not properly predict wildlife responses to climate change.

Here, we present a 131-year harvest time series of Atlantic puffins (*Fratercula arctica*, hereafter ‘puffin’) from the Westman Islands in southwest Iceland. This is the world’s largest aggregation of puffins with an estimated 830,000 breeding pairs in 2010 (Hansen et al., 2011), but has recently experienced a pronounced decline (Lilliendahl et al., 2013). Puffins have been harvested on these islands since people first settled there during the 9th century (Sveinsson, 1954), and harvest continued until 2010. Harvest records are available from 1880. As most of the puffins harvested were immature birds, we use harvest rates to calculate an index of offspring production. This provides a unique opportunity to analyse the demographic response to climate variability of a seabird over the course of more than a century, and thereby to assess the temporal consistency of such relationships (Durant, Anker-Nilssen, et al., 2004).

As apex predators in the North Atlantic, puffins respond to climate-related changes at more basal trophic levels. This means that a wide variety of other species can be expected to respond to the same climatic patterns by which puffins are affected, including many important fish stocks for which no long time series are available.

As the puffin population has been decreasing over the past century, while temperature has been increasing, we hypothesise warming to have a negative effect on offspring production overall. The time series is long enough to identify nonlinearities, however, so that one can expect to find evidence of a convex thermal performance curve. The time series covers the so-called early 20th century Arctic warming (1920s–1940s; Rogers, 1985), and we thus predict that puffin offspring production was reduced during this warming period. We further hypothesise that the demographic responses observed in recent decades, due to global warming, are similar to those observed during the early 20th century Arctic warming.

## 2 | MATERIALS AND METHODS

### 2.1 | Harvest data

The Westman Islands (Vestmannaeyjar) are an archipelago 10–30 km off the southwestern coast of Iceland (63°25′N, 20°20′W). Traditionally, puffins were hunted both for their meat and feathers. From 1876, pole nets were used and soon became the only hunting method in the Westman Islands (Árnason, 2012). Harvest took place from a limited number of locations within each colony, and these harvest locations were constant across years. Normally all suitable locations (for a given wind direction) were manned. Most harvest in a colony (>60%) was due to only one of the locations, which was manned by the most experienced hunter. These procedures entail
that the main factor affecting harvest effort was the number of harvest days during a season. Whereas the precise number of harvest days is not available for all colony-years, the length of the harvest season is. As harvest took place on all days with suitable wind conditions during the legal harvest season, the length of the harvest season is a good predictor of the number of harvest days. The time series was composed of four periods.

1. 1880–1893: Estimates were based on public feather export records (Anonymous, 1885–1894). These were converted to individual birds using an average proportion of 40 birds per kilogramme (Árnason, 2012). During this period, harvest lasted for approximately 70 days per season.

2. 1898–1941: Estimates were based on public harvest records (Anonymous, 1899–1913, 1914–1944). These records systematically underestimated the true harvest by the practice of subtracting expenses of landowner share and transport costs, both a fixed proportion of the harvest (Árnason, 2012). Good documentation of this practice exists for Elliðaey Island during 1910–1946, providing an annual multiplication factor of 1.464 ± 0.011 (see Supporting Data S1). Until 1917, a harvest season of 30 days was assumed (5 weeks, with harvest banned in certain cliffs on Heimaey that together represent 13.2% of the harvest). From 1918 and onwards, harvest lasted for c. 45 days per season.

3. 1944–1968: Estimates were based on unpublished records of the Álsey Island hunting club. The total annual harvest was estimated by dividing the Álsey harvest with Álsey’s mean annual proportion of the total harvest during the fourth period (see Supporting Data S1).

4. 1969–2010: Estimates were based on unpublished records from eight hunting clubs. Missing data (91 of 400 club years, 10.5% of the total harvest number) were interpolated using Trim (Pannekoek & van Strien, 2001; ter Braak et al., 1994), except for the Heimaey Island common. For the latter, only 5 years of harvest are known (1947–1951), and, based on those figures, harvest was assumed to be a fixed proportion (1.566 ± 0.155) of the harvest in the Ystiklettur hunting club (see Supporting Data S1). The hunting season was incrementally reduced from 45 days to 30 days in 2008, to 20 days in 2009 and to 5 days in 2010. Hunting has been virtually banned since 2011.

The harvest time series thus obtained had six missing values (1894–97 and 1942/43), which were filled in using cubic spline interpolation. The differing lengths of harvesting seasons were corrected for by standardizing all data to a season of 45 days (assuming approximately constant daily harvest).

2.2 | Chick production index

Chick production \( P_t \) in a given year \( t \) was estimated as

\[
P_t = C \cdot \sum_{a=2}^{6} p_a \cdot H_{t+a} \cdot \phi^{-a},
\]

where \( C \) is a constant scaling factor; \( p_a \) is the proportion of harvested birds in age class \( a \); \( H_t \) is the harvest in year \( t \); and \( \phi \) is the annual adult survival rate.

The variables \( p_a \) and \( \phi \) were inferred as follows: from 1959 onwards, puffin chicks have been ringed at fledging with stainless steel rings. Hunters collected and returned rings of harvested birds. Based on these data \( N = 7662 \), the age distribution of harvested ringed birds could be inferred (Helgason, 2012; unpubl. data). 82% of the harvested birds were between 2 and 6 years old. The remaining age classes (1 year and ≥7 years) were ignored in the analyses, because they occurred at very low frequencies per cohort (<2.5%; Table S1). The proportions \( p_a \) were thus 0.199 ± 0.003 (a = 2 years), 0.416 ± 0.004 (a = 3 years), 0.253 ± 0.002 (a = 4 years), 0.098 ± 0.001 (a = 5 years) and 0.034 ± 0.001 (a = 6 years). These proportions are consistent with the age-related changes in colony attendance (Sandvik et al., 2008), since puffins start prospecting their breeding colony at 2 years of age and start breeding at around 5 years of age (Ashcroft, 1979; Breton et al., 2006; Harris & Wanless, 2011). Prospecting (pre-breeding) individuals spend much time flying around the colony and are thus overrepresented among the harvested birds. After an individual has become a breeder, it spends most time on the nest (in breeding burrows) and foraging (at sea), so that the harvest rate is lower for older age classes.

Equation 1 corrects for mortality by dividing the proportion \( p_a \) of harvested birds of age \( a \) by the cumulative survival \( \phi^a \) until that age (where \( \phi \) is estimated from the ring recoveries to be 87.2%; Table S1; cf. Helgason, 2012). Finally, all sums were multiplied by \( C = 22,147^{-1} \) so as to scale chick production indices between 1 and 10 (Figure 1).

Chick production is a function of population size, and we therefore log-transformed the chick production index prior to analysis. For ease of interpretation, we present the untransformed (or back-transformed) index in graphs. On the untransformed scale, an index of 0 thus corresponds to no chick production.

It should be emphasised that the chick production index is a proxy for the relative cohort size fledged in a given year. First, no attempt is made to infer the absolute cohort size, because annual population estimates do not exist for most years of the 128-year period. Second, the estimation does not correct for variation in harvest effort, beyond the corrections described in the section Harvest Data. Consequently, the chick production index must be regarded as a smoothed time series, as it underestimates the year-to-year variation, although it should be expected to be a good proxy of the mid- to long-term variation (above c. 5 years) and free from systematic bias. By back-estimating the yearly harvest based on the chick production index (by the reversed procedure outlined above), one can infer the signal-to-noise ratio of the proxy; the coefficient of determination between the reported harvest and the back-estimated harvest time series is \( r^2 = 0.821 \).

Analyses were based on the chick production indices from 1878 to 2005 (128 years; Supporting Data S1). To ensure that our underlying assumptions did not create spurious results, analyses were re-run for four alternative datasets: (a) a shorter time series (1896–2000, i.e. 105 years), omitting uncertain estimates in the beginning
and the change in the harvesting regime at the end; (b) a series based on linear interpolation of the six missing values; (c) a series that was uncorrected for the length of the harvest season; (d) the harvest series itself (1881–2008), lagged by 3 years (which was the median and modal age of harvested birds).

### 2.3 Environmental covariates

To explore the relationship between the puffin time series and temperature, we used local data on sea surface temperature (SST). Monthly mean SST and air temperature data from Heimaey Island, which is the main island of the Westman archipelago, were supplied by the Icelandic Meteorological Office and the Marine and Freshwater Research Institute. A gap in SST data between 1964 and 1998 was interpolated by regressing SST on air temperature for each month (see Table S2). SST in year \( t \) was defined as the average SST from September \( t-1 \) to August \( t \). This definition is in accordance with the breeding phenology of puffins, which stay in or close to the breeding colony from March/April until August/September, so that potential effects during the non-breeding season would only become visible in the following breeding season. In order to also represent SST on a larger spatial scale, we used the Atlantic Multidecadal Oscillation (AMO) index, a detrended time series of SSTs over the entire North Atlantic (Schlesinger & Ramankutty, 1994). We used AMO data based on the HadISST1 dataset (Smith, 2015).

Four climate indices were used as covariates. (1) The North Atlantic Oscillation (NAO) is one of the dominant teleconnection patterns in the North Atlantic (Hurrell et al., 2003). We used the annual station-based NAO index, because it is available from 1864 onwards (see Hurrell, 1995; Hurrell & National Center for Atmospheric Research Staff, 2017). (2) The Arctic Oscillation (AO) is closely related to the NAO but has a hemispheric and northerly pattern (Thompson & Wallace, 1998). We used annual means of monthly AO indices from the 20th Century Reanalysis (Smith, 2013). (3) The North Atlantic subpolar gyre (SPG) is located in the northern North Atlantic, 3000 km southwest of Iceland, and dominates the physical oceanography of this region (Hátún et al., 2005). The SPG index series was provided by Helge Drange (Univ. of Bergen, pers. comm.). (4) The El Niño-Southern Oscillation (ENSO) is a climatic variation pattern over the tropical eastern Pacific (Trenberth & Stepaniak, 2001). We represented ENSO by the Niño 3.4 SST index (Smith, 2018).

In other North Atlantic systems, seabird demography has been shown to covary with climate indices such as AO (Smith & Gaston, 2012), NAO (Durant, Stenseth, et al., 2004; Reid et al., 1999; Sandvik et al., 2005) and SPG (Descamps et al., 2013; Fluhr et al., 2017; Hátún et al., 2017). Given that strong El Niño events can affect conditions on the northern hemisphere (Brönnimann et al., 2004, 2007), one might even expect an ENSO-related signal.

### 2.4 Statistical analyses

Linear and quadratic responses to the covariates were investigated using linear models. In order to obtain unbiased error estimates, quadratic responses were estimated using nonlinear least squares models with the following model structure:

\[
\ln P_t \sim \ln P_{\text{peak}} + b \cdot (C_t - C_{\text{peak}})^2, \tag{2}
\]

where \( P_{\text{peak}} \) is the peak of the chick production index, \( b \) is the slope of the quadratic term, \( C_t \) is the value of the covariate in year \( t \), and \( C_{\text{peak}} \) is the value of the covariate at which \( P_t \) peaks. For the case of a change in \( C_{\text{peak}} \) over time, Equation 2 was modified to:

\[
\ln P_t \sim \ln P_{\text{peak}} + b \cdot (C_t - C_{2000} - b_{\text{peak}} (t - 2000))^2, \tag{3}
\]

where \( C_{2000} \) is the value of the covariate at which \( P_t \) would have peaked in the year 2000, and \( b_{\text{peak}} \) is the slope with which \( C_{\text{peak}} \) changes.
Confidence intervals of nonlinear models were estimated using R package ‘invest’ (Greenwell & Schubert Kabban, 2014).

Model selection was based on Akaike’s information criterion (AIC; Burnham & Anderson, 2002). Our data violated one essential assumption of standard linear models: adjacent data points are not statistically independent, because there is strong temporal autocorrelation in the time series. This does not affect the variance explained, but it invalidates the p values estimated using standard tests as well as inflating AIC values. We therefore estimated p values based on a parametric bootstrap procedure (Manly, 1997), i.e. we created 100,000 time series with the same length and autoregression-moving average (ARMA) structure as the harvest data, namely ARMA(p = 1, q = 0). Equation 1 was then used to calculate chick production indices from the simulated harvest series. A p-value, referred to as $p_{boot}$ was calculated as the proportion of time series that obtained an $r^2$ equal to or larger than the one obtained with the real chick production index.

A good model fit may be obtained when a nuisance parameter affecting puffin harvest (e.g. economic or societal changes) incidentally has a parallel long-term pattern to one of the covariates. In order to correct for this possibility, we carried out a second bootstrap test, in which we removed the long-term pattern (regression splines with 4 df) of each of the 100,000 simulated time series, and added the actual long-term pattern of the chick production index. This procedure guaranteed that the 100,000 simulations differed from the actual time series only in their medium- and short-term dynamics, so that a good model fit of the real data cannot be due to the long-term pattern. The p-value obtained, referred to as $p_{long}$ thus corrects for potential long-term changes in nuisance parameters.

As a post-hoc test of the sensitivity to potential biases when combining harvest data from four different subperiods (see Section 2.1 above), we considered the possibilities that each of the first three subsets of the time series should be up- or downweighted by 10% or 20% (i.e. multiplied by 0.8, 0.9, 1.1 or 1.2) relative to the fourth subset. Based on these four different weightings for the three subsets, we calculated $4^3 = 64$ time series that varied in weighting factors. The two best models identified in previous analyses were then retested for each of these time series using a parametric bootstrap with 10,000 simulations (as described for the other datasets).

Models were also investigated at different time scales, viz. long-, medium- and short-term, by filtering away variability at other time scales. The long-term (appr. centennial) pattern was defined as regression splines with 4 df fitted in a generalized additive model. The medium-term (appr. decennial) pattern was defined as a time series smoothed using 11-year running means. The short-term (year-to-year) pattern was defined as the residuals from the long-term pattern. The temporal structure of variables was investigated using generalised additive models and wavelet analysis, using R packages ‘nlme’ and ‘WaveletComp’ (Pinheiro et al., 2020; Roesch & Schmidbauer, 2018).

Changes over time in temporal trends of, and correlations between, chick production and covariates were analysed using sliding windows. Sliding windows were centred on the year of interest and were 11 and 31 years long for trends and correlations, respectively (one or three decades, plus 1 year to obtain symmetrical intervals on each side of the year of interest).

Unless stated otherwise, calculations were carried out in R (R Core Team, 2020). Estimates are reported as mean ± standard error (SE).

## 3 | RESULTS

Puffin harvest on the Westman Islands reached an early maximum of 175,000 birds in 1885, whereupon it declined to 14,000 harvested birds in 1946 (Figure 1). Harvest numbers increased again, reaching 143,000 birds in 1998, and dropped quickly thereafter. In 2007, the last year without hunting restrictions, 31,500 birds were harvested.

The chick production index closely tracks the temporal fluctuations in harvest, from which it is estimated. A considerable amount of the variance in chick production was explained by SST, whereas none of the climate indices explained more than 7% of the variance, with AMO being intermediate (Table 1). For SST, but not the other covariates, quadratic models received more support than the linear counterparts. Smoothed SST (11-year running means) performed even better than annual SST values, explaining almost 72% of the variance in chick production. Corresponding results were found with the modified datasets (Tables S3–S7).

According to the best model among the set of models compared (Table 1), chick production reached a peak index value of 6.6 ± 0.2 at an SST of 7.09°C ± 0.05°C (Figure 2; Table S8). At SSTs of 1°C above or below this peak, chick production was reduced by 55% ± 2%.

### 3.1 | Patterns identified at different temporal scales

Patterns of covariation between puffin chick production were further investigated at three temporal scales, viz. long-term (approx. centennial), medium-term (approx. decennial) and short-term (year-to-year variability). The long-term pattern in chick production followed the long-term pattern in SST very closely (Figure 3). This pattern could be confirmed using cross-wavelet power spectra, which indicated a common signal for chick production and SST with a period of 75–100 years (peaking at 84 years; Figure S1).

On a medium temporal scale, the co-variation in chick production and SST exhibited strikingly variable patterns. First, the correlation between chick production and annual mean SSTs was not negative over the entire period, but only during the sub-period between 1920 and 1970, and again towards the end of the time series (Figure 4). Prior to, and in between these sub-periods, the correlation switched to a positive sign. This temporal pattern was even more apparent when comparing decennial trends in both variables (Figure 5): as indicated by the vertical lines in Figures 3 and 5, the decennial trends of production and SST were in phase exactly when the long-term SST...
was below its long-term mean, and in antiphase exactly when SST was above its long-term mean. This pattern is highly unlikely to emerge by chance (three randomly placed vertical lines produced an equally strong or stronger relationship in only 0.02% of the possible cases).

The sub-periods in Figures 3 and 5 were used to define an additional categorical variable (warm vs. cool phase). Models including SST, sub-period and their interaction were intermediate to the corresponding linear and quadratic ones (annual SST, $r^2 = 0.55$, $\Delta$AIC = 61.9; smoothed SST, $r^2 = 0.68$, $\Delta$AIC = 18.5). All these findings reinforced the suggestion that, as far as puffin reproduction is concerned, under warm conditions, ‘cooler is better’, whereas, under cool conditions, ‘warmer is better’.

Here, ‘cool’ and ‘warm’ are relative to the peak at 7.1°C. There was some evidence that the peak temperature itself was not constant through time, but has rather been increasing slightly during the 20th century. A model in which the peak SST was allowed to change linearly was somewhat better supported than the best model of Table 1 ($r^2 = 0.73$, $\Delta$AIC = -5.2). It indicated that the peak was at 7.27°C ± 0.08°C in the year 2000, which was 0.24°C ± 0.09°C warmer than in the year 1900. This finding was derived from the smoothed SST data and could not be reproduced with the annual SST data and only with some of the alternative datasets (Table S9).

The short-term (year-to-year) variability in chick production was not captured well by any of the explanatory variables. Residual SST explained less than 1% of the variation in residual chick production. Even in a multiplicative model with sub-period and SST, the explained variance remained below 0.1 ($p_{\text{boot}} = 0.063$).

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Model structure</th>
<th>$\Delta$AIC</th>
<th>$r^2$</th>
<th>Estimate ± SE</th>
<th>$p_{\text{boot}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoothed SST</td>
<td>quadratic</td>
<td>0.00</td>
<td>0.717</td>
<td>-0.792 ± 0.073</td>
<td>0.00005</td>
</tr>
<tr>
<td>Annual SST</td>
<td>quadratic</td>
<td>68.36</td>
<td>0.517</td>
<td>-0.383 ± 0.056</td>
<td>0.00054</td>
</tr>
<tr>
<td>Smoothed SST</td>
<td>linear</td>
<td>83.50</td>
<td>0.447</td>
<td>-0.665 ± 0.066</td>
<td>0.011</td>
</tr>
<tr>
<td>Annual SST</td>
<td>linear</td>
<td>106.84</td>
<td>0.337</td>
<td>-0.506 ± 0.063</td>
<td>0.015</td>
</tr>
<tr>
<td>AMO</td>
<td>linear</td>
<td>113.97</td>
<td>0.299</td>
<td>-1.56 ± 0.21</td>
<td>0.0052</td>
</tr>
<tr>
<td>AMO</td>
<td>quadratic</td>
<td>115.49</td>
<td>0.301</td>
<td>-0.74 ± 1.08</td>
<td>0.0067</td>
</tr>
<tr>
<td>AO</td>
<td>linear</td>
<td>151.15</td>
<td>0.063</td>
<td>0.34 ± 0.12</td>
<td>0.054</td>
</tr>
<tr>
<td>AO</td>
<td>quadratic</td>
<td>152.82</td>
<td>0.065</td>
<td>0.11 ± 0.20</td>
<td>0.082</td>
</tr>
<tr>
<td>NAO</td>
<td>linear</td>
<td>157.25</td>
<td>0.017</td>
<td>0.036 ± 0.024</td>
<td>0.20</td>
</tr>
<tr>
<td>Null model</td>
<td>–</td>
<td>157.41</td>
<td>0.000</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>SPG</td>
<td>linear</td>
<td>158.30</td>
<td>0.009</td>
<td>0.022 ± 0.021</td>
<td>0.73</td>
</tr>
<tr>
<td>NAO</td>
<td>quadratic</td>
<td>158.50</td>
<td>0.023</td>
<td>0.011 ± 0.013</td>
<td>0.32</td>
</tr>
<tr>
<td>ENSO</td>
<td>quadratic</td>
<td>158.53</td>
<td>0.022</td>
<td>0.138 ± 0.096</td>
<td>0.21</td>
</tr>
<tr>
<td>ENSO</td>
<td>linear</td>
<td>158.61</td>
<td>0.006</td>
<td>-0.068 ± 0.076</td>
<td>0.32</td>
</tr>
<tr>
<td>SPG</td>
<td>quadratic</td>
<td>160.30</td>
<td>0.009</td>
<td>0.000 ± 0.008</td>
<td>0.82</td>
</tr>
</tbody>
</table>

**TABLE 1** Models explaining temporal variation in puffin chick production on the Westman Islands (Iceland), 1878–2005, using different environmental covariates. Models are sorted by decreasing fit (increasing $\Delta$AIC). The AIC of the best model was 39.27. For quadratic models, the estimate of the slope b is provided (see Equation 2). Probabilities ($p_{\text{boot}}$) were derived by bootstrapping (see Section 2).
DISCUSSION

Based on 131 years of harvest data, we identified prolonged periods of strong correlations between sea surface temperature (SST) and productivity in Icelandic puffins. The sign of the correlation switched several times, being strongly negative from the 1920s to the 1970s, and positive before and after this 50-year period (Figure 4). This pattern is well explained by a quadratic relationship between SST and chick production, which peaks at an SST of 7.1°C. In years that are warmer or cooler than the peak SST, chick production is lower (Figure 2). Because SST changed rather abruptly in 1921 (upwards) and 1967 (downwards) and increased steeply also in the late 1990s (Figure 3), most of the variation during the sub-periods is on one side of the peak SST, thus resulting in prolonged periods of positive or negative correlations.

Such rapid changes in the marine environment have been reported in the Atlantic in different time periods, some being described as regime shifts. Two of the abrupt changes in the sign of the correlation between puffin chick production and SST (Figure 4) coincided with large-scale shifts occurring quasi-synchronously over the entire Northern Hemisphere (1975/76 and 1988/89; Beaugrand et al., 2015). These regime shifts were associated with large changes in plankton and fish communities and/or abundance (Beaugrand & Reid, 2012), which could potentially explain the effect on puffin chick production through bottom-up effects.

The 50-year period with a strong negative correlation between sea surface temperature and chick production (Figures 4 and 5) concurred with the so-called early 20th century Arctic warming (Grant et al., 2009; Rogers, 1985; Scherhag, 1939; Yamanouchi, 2011). The onset of wintertime warming, estimated to be 1921 for Iceland (Rogers, 1985), agrees well with the first sign shift in our Figure 5.
The causes and attribution of the early 20th century Arctic warming remain uncertain; while there is a contribution from anthropogenic forcing by greenhouse gases (Abram et al., 2016; Callendar, 1938; Schurer et al., 2013), the contribution of natural decadal variability in the climate system (internal forcing) seems to be more important (Grant et al., 2009; Yamanouchi, 2011). The congruence between changes in the sign of the correlation between SST and chick production, and the occurrence of large-scale and abrupt changes in the marine environment confirm the potential importance of the rate of environmental changes and in particular of warming (Descamps et al., 2017; Irons et al., 2008).

It is a strength of long time series that such non-linear patterns can be revealed. Based on a shorter time series, one might easily have ended up concluding differently, viz. that the relationship was uniformly negative or positive (or absent). For the identification of a variable peak SST, time series need to be even longer. We also found some evidence for non-stationarity, that is, that the temperature at which production peaks has been increasing by roughly a quarter of a degree during the 20th century, although this increase was weak and could not be corroborated with all modified datasets (Table S9).

The observed change in the relationship between SST and puffin productivity has huge implications for our understanding of climate warming consequences on wildlife. Indeed, the vast majority of studies quantifying the relationships between wildlife and changes in air or sea temperature is based on short time-scale, from just a few years (e.g. Hovinen et al., 2014) to a few decades (e.g. Barbraud & Weimerskirch, 2001; Price et al., 2020). Many of these studies assume a linear and stationary effect of the environment, and so do the predictions about the fate of populations (but see e.g. Barbraud et al., 2011; Descamps et al., 2017; Jenouvrier et al., 2012). Our study clearly shows that this may be erroneous, as the effect of a warming environment may vary through time and easily be hidden by shorter-term variation.

Nestling growth rate and fledging success in a close relative, the tufted puffin (F. cirrhata), have also been shown to vary around a peak SST (Gjerdrum et al., 2003). Some studies, compensating for short time series by combining data from several colonies, have demonstrated similar concave patterns in seabird demography (Grosbois et al., 2009; Iles et al., 2020). Interestingly, one such study has shown adult survival of puffin to peak at a spring SST of c. 7°C (Grosbois et al., 2009).

Based on the feeding ecology of Icelandic puffins, we hypothesize that the biology of their main prey species, the lesser sandeel (Ammodytes marinus), might be the key to explaining the temporal dynamics. Sandeels constitute the predominant prey of puffins in the Westman Islands (Lilliendahl et al., 2013; Lilliendahl & Solmundsson, 1997). Bioenergetic models have shown that 0-group sandeels need to reach a critical size threshold before their first winter, in order to survive and reach maturity; and that the critical size threshold increases with increasing temperature (van Deurs et al., 2011). Furthermore, sandeels develop smaller gonads under higher winter temperatures (Wright et al., 2017).

In accordance with these findings, sandeel recruitment in the North Sea was reduced under conditions with warmer sea temperatures (Arnott & Ruxton, 2002); and the body size of both 0-group and 1-group fish has decreased in recent decades (Wanless et al., 2004, 2018). Due to higher mortality and lower fecundity of sandeels, puffins may thus face food scarcity after relatively warm winters. However, sandeel presence and abundance also decrease at lower temperatures, with a bottom temperature of 8.5°C representing the optimum for sandeel on the Dogger Bank (van der Kooij et al., 2008). This is higher than the peak at 7.1°C identified in this study and may reflect differences in the local adaptations of sandeel between the two areas; such adaptations are known from other species (Conover & Present, 1990).

The trophic explanation is strengthened by the fact that Icelandic puffins do not overwinter close to the colonies (but as far west as the Labrador Sea; Fayet et al., 2017), whereas the local winter SST off the Westman Islands is a better predictor of chick production than AMO, which represents SST at a larger spatial

![Figure 5](https://wileyonlinelibrary.com)
scale. This indicates that the effect of winter temperature on puffins is unlikely to be direct (i.e. on adult overwinter survival or condition), but must be assumed to be mediated via the food web (cf. Sandvik et al., 2012).

If the conditions during the early 20th century Arctic warming can be extrapolated to the future, the projected warming of the oceans will have negative effects on puffin offspring production in southwest Iceland, a present and historical stronghold for this species. Figure 4 indicates that the years after 1989 might indeed have been the beginning of such a period. The consequence would be a decline in the Icelandic puffin population, unless the population response is buffered by an opposite response in other demographic parameters (especially adult survival).

The increase in the temperature at which the puffins’ production peaked, provided that this finding is not an artefact, may partially compensate for the warming sea water. However, the increase in average SST during the 20th century has been more than three times faster (0.79°C ± 0.12°C) than the concurrent increase in the peak SST (0.24°C ± 0.09°C). Due to the long generation time of puffins (11 years; Bird et al., 2020), evolutionary adaptation in puffins is unlikely as an explanation for this trend. More likely explanations are adaptations in the prey (sandeels) or plastic behavioural responses by the puffins (e.g. partial switching to other prey sources).

Direct productivity measurements show that the population is already declining; indeed, the very reason for banning puffin harvest in 2011, was a concern for population persistence (Lilliendahl et al., 2013). Based on the models presented here and the historical parallel during the 1940s, it is very likely that this decline is at least partially caused by the increasing SSTs around Iceland, and that the situation will continue to aggravate with global warming. Our study also highlights the challenges associated with documenting bottom-up effects of climate change in large and highly variable marine ecosystems and the value of centennial time series in this context: Unexpected and surprising reversals of decennial trends may become understandable when the underlying patterns are better understood.

ACKNOWLEDGEMENTS
We thank the seabird hunting clubs of the Westman Islands for access to their harvest data. Our work was supported by funding from the Fram Centre (High North Research Centre for Climate and the Environment) by a grant to K.E.E. (Flagship in Coastal Ecology), H.S. was also supported by the Research Council of Norway through its Centres of Excellence funding scheme (project number 223257). Kate Layton-Matthews commented on the manuscript.

CONFLICT OF INTEREST
The authors state that they have no competing interests.

AUTHOR CONTRIBUTIONS
All authors participated in designing the study and in interpreting the results. Data on puffin harvest were assembled by E.S.H. Analyses were performed by H.S. with input from K.E.E. and N.G.Y. The paper was written by H.S. with input from all co-authors.

DATA AVAILABILITY STATEMENT
The data analysed and the code used during this study are included as Supplementary Data and Supplementary Information accompanying this article.

ORCID
Erpur S. Hansen https://orcid.org/0000-0001-6899-2817
Hanno Sandvik https://orcid.org/0000-0002-5889-1606
Kjell Einar Erikstad https://orcid.org/0000-0002-2171-5721
Nigel G. Yoccoz https://orcid.org/0000-0003-2192-1039
Tycho Anker-Nilsen https://orcid.org/0000-0002-1030-5524
Jürgen Bader https://orcid.org/0000-0001-9294-3169
Sébastien Descamps https://orcid.org/0000-0003-0590-9013
Kevin Hodges https://orcid.org/0000-0003-0894-229X
Michel d. S. Mesquita https://orcid.org/0000-0002-4556-5414
Tone K. Reiertsen https://orcid.org/0000-0002-9579-2420
Øystein Varpe https://orcid.org/0000-0002-5895-6983

REFERENCES
Anonymous. (1914–1944). Dúntekja og fuglateki [Feather and bird harvest (in Icelandic)].


Bird numbers in the wintertime geopotential height and temperature variation in the Northern Hemisphere. Marine Ecology Progress Series, 454, 237-249. https://doi.org/10.3354/meps110826.1

Climate change and marine vertebrates. Science, 350, 772-777. https://doi.org/10.1126/science.aac9874


SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section.