Varying population size of the Cape Royds Adélie penguin colony, 1955-2020: a synthesis

Supplementary Material

Details of ARIMA Modeling

Annual growth data were analyzed using the Auto Regressive Integrated Moving Average (ARIMA) function, which is widely used in the forecasting of time series data (Box *et al.* 2015). The fitted ARIMA model was used to assess the stationarity (constancy of variance), fluctuation, and autocorrelation of the data. The ARIMA model included two modules, Autoregression (AR) and Moving Average (MA), which can be integrated to describe statistical patterns when the data are not stationary but differenced to achieve stationarity.

The ARIMA model is represented as ARIMA (*p, d, q*), where *p* indicates the order of the autoregression, *d* indicates the order of the difference, and *q* indicates the order of the smoothing moving average. The AR (*p*) process assumes that the recent value can be predicted from previous (lagged) values, generally:

$$y_t = \phi_{t-1} y_{t-1} + \phi_{t-2} y_{t-2} + \cdots + \phi_{t-p} y_{t-p} \quad (1)$$

where y_t is the value of the response variable at time *t*, y_{t-1} is the value of the response variable at time *t*-1thus lagging 1 time period, and Φ_1 is the coefficient of the effect of the response variable at time *t*-1, thus being the first autoregressive coefficient. The difference component of an ARIMA model refers to accounting for seasonality in the data (*d*). Because we use yearly data, the use of this parameter would be warranted in the presence of multi-year cycles. Multi-annual cycles were immediately apparent, and we explored models that would include them. The MA

(q) adds a smoothing of the error at each time lag, derived from the previous time period. Thus, the ARIMA (1,0,1) model is:

$$y_t = \phi_{t-1} y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad (2)$$

where ε_t is the immediate error, ε_{t-1} is the error at time t-1, and θ_1 is the smoothing coefficient for the error component at $t-1$.

We sought to explain the variance in the current observation (year) after accounting for the autocorrelation or cycle components with our explanatory variables and, therefore, the use of the MA component was not warranted. Thus, we fitted some variations including only structures where q was zero in the ARIMA models: ARIMA(1,0,0), ARIMA(2,0,0), ARIMA(3,0,0), ARIMA(1,1,0), ARIMA(2,1,0), ARIMA(3,1,0). The appropriate lag (p) and differencing (d) for each species and colony were assessed by examining the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots.

Finally, we checked the residuals from our chosen model by plotting the ACF of the residuals (Table SM1). To demonstrate covariate effects, we generated variable influence plots as follows, obtaining predictions from the models and their 95% confidence intervals. In the absence of any other covariates, we could have simply regressed the predictions against the values of the covariate. However, the predictions include the effects of all other covariates as well and thus such a regression would not truly portray the effect of each variable alone in the model. To overcome this limitation, we first regressed the variable of interest against all other variables and extracted the residuals of this regression. We then added these residuals to the mean value of the variable of interest, thereby obtaining the values of the variable not explained by collinearity with the other variables. We then plotted predictions against this adjusted covariate.

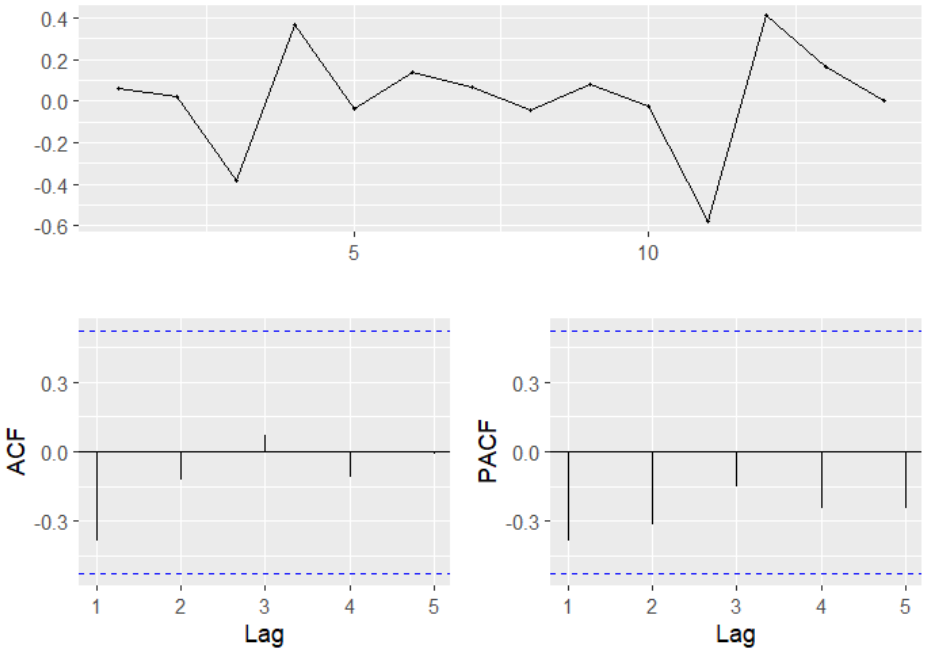
Table SM1. Different percentage weights (in tonnes extracted) explored for "fish cum" (cumulative fish extracted).

Fish = Total Tonnes taken * Percent ≥134 TL/100

Fish_cum 3	3 years cumulative, no weights
Fish_cum_8_9_1	3 years cumulative, weights: 80%, 90%, 100%
Fish_cum_6_8_1	3 years cumulative, weights: 60%, 80%, 100%
Fish_cum_75_1	2 years cumulative, weights: 75%, 100%
Fish_cum 2	2 years cumulative, no weights
Fish_cum_5_1	2 years cumulative, weights: 50%, 100%
Fish_cum_4_1	2 years cumulative, weights: 40%, 100%
Fish_cum_2_1	2 years cumulative, weights: 20%, 100%

Figure SM1. Results from first order autocorrelation ARIMA modeling, including ACF (Autocorrelation Function) and PACF (Partial Autocorrelation Function), to evaluate annual growth trend of the Cape Royds Adélie penguin colony.

1.1) ARIMA (1,0,0) Uncorrected trend



1.2) ARIMA (1,0,0) Residuals with drift