Age estimation of North Atlantic Right Whales by allometric measurements on photographs

Mariano Sironi[1,2], Scott D. Kraus[3], Erik V. Nordheim[4], Victoria J. Rowntree[5] and Charles T. Snowdon[1]

[1]Departments of Zoology and Psychology, University of Wisconsin, Madison, USA

[2]Instituto de Conservación de Ballenas, Buenos Aires, Argentina

[3]New England Aquarium. Central Wharf. Boston MA 02110, USA

[4]Department of Statistics and Department of Forest Ecology & Management, University of Wisconsin, Madison, USA

[5]Whale Conservation Institute, Department of Biology, University of Utah, Salt Lake City, USA

Contact e-mail: msironi@icb.org.ar

ABSTRACT

Models to predict the age of North Atlantic right whales (*Eubalaena glacialis*) were developed using shipboard lateral photographs of their heads. Models based on allometric ratios of head measurements explained three quarters or more of the variability in the square root of age as response variable. Model fit was lower for whales with continuous callosity patterns. Changes in the curvature of the rostrum and the height of the dome were the best predictors of age in the external anatomy of the head of right whales. The accuracy of age prediction decreased with age: it was maximum for calves and minimum for whales older than 8yr. Mean coefficients of variation of repeated measurements ranged from 0.31 to 4.11%. Photographs taken at medium distance from the whales were better than those taken close or far, to avoid biases related to the angle of view of the subjects and to the small size of the measurements, respectively. This photogrammetric technique is non-intrusive, simple and inexpensive. Because it is based on measurement ratios, it eliminates the need to know the distance to the subjects or to have a scale object close to them to estimate the age of live right whales.

KEYWORDS: AGE DETERMINATION, NORTH ATLANTIC RIGHT WHALE, MODELLING, PHOTOGRAMMETRY, REGRESSION

INTRODUCTION

Knowing the age of individuals in the wild is critical for understanding population dynamics and for estimating the demographic parameters of a species. Depleted whale stocks must be monitored by means of benign techniques that do not rely on disturbing or killing the whales to determine life history parameters and population trends (Whitehead and Payne, 1981; Kraus *et al.,* 1986; Best and Rüther, 1992; Spitz *et al.,* 2000).

In this context, several non-invasive techniques have been used to measure the size and thus estimate the age of cetaceans. Some rely on a scaling object of known size that is photographed from an aircraft in the same frame as the whale that is being measured (Whitehead and Payne, 1981). Others combine altimeter readings and the focal length of the camera lens to estimate the size of whales from aerial photographs (Cubbage and Calambokidis, 1987; Best and Rüther, 1992; Ratnaswamy and Winn, 1993; Angliss *et al.,* 1995; Perryman and Lynn, 2002). Photographs taken from boats also allow estimation of body sizes of sperm whales (Gordon, 1990; Dawson *et al.,* 1995). Underwater videogrammetry has been used to measure humpback whales while gathering data on individual identification and behavior (Spitz *et al.* 2000).

Right whales (*Eubalaena*) can be identified individually by the unique pattern of callosities on their heads (Payne *et al.,* 1983; Kraus *et al.,* 1986). Long-term studies have produced photoidentification catalogs of southern right whales (*E. australis*) in the SW Atlantic (Payne and Rowntree, 1984) and North Atlantic right whales (*E. glacialis*) in the NW Atlantic (Hamilton and Martin, 1999). Photographs are taken from a variety of platforms, including boats and airplanes. The small size of the callosity pattern of newborn calves and the large patches of cyamids on their heads (Rowntree, 1996) make the identification of calves difficult. As a consequence, whales that were not identified in their year of birth and that are resighted in later years are of unknown age.

The head of right whales shows positive allometric growth from their calf year to adulthood, although the relationship is not sufficiently strong to estimate growth during the first year of life (Whitehead and Payne, 1981; Best and Rüther, 1992). The overall shape of the head changes with age; in profile the head appears "truncated" and with a raised coaming in calves (Hamilton and Martin, 1999). The rostrum flattens as the animals grow. These head shape changes, if properly quantified, could provide a basis for age estimation.

We developed a technique to estimate the age of N Atlantic right whales based on a combination of ratios of allometric measurements of their head taken from lateral shipboard photographs of whales of known ages. The main aim of the technique is to predict the age of whales whose year of birth is unknown using allometric ratios. The application of this method could allow better estimations of population parameters that are based on the age of individuals. The technique could be applied to all right whale populations and could improve the assessment of the demographic parameters of right whales throughout the northern and southern hemispheres.

MATERIALS AND METHODS

Photographs used

Measurements were taken on 374 photographs of 62 N Atlantic right whales (30 females and 32 males) born between 1986 and 1996. The year of birth was known for all the whales selected. The photographs are stored in the catalog compiled at the New England Aquarium in Boston (for a catalog description, see Hamilton and Martin, 1999). The photographs used were shipboard color slides that showed lateral views of the rostrum, the upper portion of the lower lip, and the back of the whales above the water surface. Photographs of some whales were not available for some years due to individual differences in sighting and photographic records. A description of photoidentification problems of N Atlantic right whales can be found in Kraus *et al*. (1986).

Measurements taken

Fig. 1 shows the measurements taken on the photographs. In order to make this technique simple to use and available to most right whale researchers, all measurements were made directly on the projected image of slides on a white paper screen to a precision of 0.5mm. The distance between the lens of the projector and the screen was 180cm, and the size of the projected image was 40 x 60cm. Sketches of the head profile (including the outline of the callosities and the lower lip) of some whales were made by projecting the image on a sheet of paper and by drawing on the paper along the profile. The sketches of each individual were visualized simultaneously to compare the changes in the head profile at different ages.

The callosity nomenclature used here follows that proposed by Payne *et al.* (1983). Some terms used in this paper are defined in Table 1. In other studies (*e.g.*, Whitehead and Payne, 1981; Best and Rüther, 1992; Ratnaswamy and Winn, 1993), the rostral length was measured from the tip of the snout to the center of the blowholes. Here, the rostral length was measured from the front tip of the bonnet (visible in more lateral photographs than the tip of the snout, which is frequently underwater) to the rear tip of the coaming (which in lateral photographs can be located with a smaller margin of error than the center of the blowholes). The height of the dome (the raised portion of the rostrum where the coaming sits) was measured from the highest point of the coaming to the imaginary line that connects the isthmus to the lower margin of the post-blowhole island. The remaining measurements are self-explanatory from Fig. 1.

Photograph quality

Two variables in the photographs (angle and endpoints) were considered as features for grading the quality of the photographs for the measurements, with 1 indicating highest quality and 4 indicating lowest quality. The position of the whales on the photographs (or "angle") was graded as 1 (whale perpendicular to the line connecting it to the photographer and horizontal at the surface, producing a full lateral view of the emerged portion of the head), 2 (slight angle: whale in slight diagonal to the photographer, *e.g.*, approaching or leaving, but flat relative to the surface), 3 (same as 2 with a marked angle) or 4 (same as 2 or 3, but whale not horizontal). In photographs graded as 4, the whale was tilted along its longitudinal axis, or the photograph was taken from a high perspective, *i.e.*, the whale was very close to the photographer standing relatively high on a boat, and the whale was viewed from above.

The endpoints of the dimensions measured on the whales' head were graded as 1 (both endpoints visible and clear), 2 (both endpoints visible but slightly blurred), 3 (one endpoint not visible but visual extrapolation was possible) or 4 (both endpoints not visible and extrapolation was necessary for both). Especially important for measuring rostrum length were the front tip of the bonnet (underwater in some photographs) and the rear tip of the coaming.

Not all measurements were affected equally by the variations in the quality of the angle and the endpoints. For example, a marked tilt along the longitudinal axis of a whale significantly affected vertical measurements (*e.g.*, valley-zenith) but did not affect longitudinal measurements (*e.g.*, rostral length) if the whale was perpendicular to the photographer. As a result, the same photograph could be graded as 1 for a variable affecting some measurements and as 4 for a variable affecting others. Measurements taken on photographs graded 1 and 2 were classified as "high quality" (HQ) and those graded 3 and 4 were classified as "low quality" (LQ). The photographs with LQ measurements were subdivided into three groups: group 1 (those with grade 3 or 4 for angle), group 2 (grade 3 or 4 for endpoints) and group 3 (grade 3 or 4 for both angle and endpoints) (see Table 2 for sample sizes). HQ and LQ photographs were analyzed separately.

A subset of photographs of whales with continuous callosity pattern was analyzed separately in order to determine how the particular features of this type of pattern may affect age prediction using the present technique.

The distance between the photographers and the whales was unknown. However, the relative size of the head of the whales on the projected image was used as an indication of distance. Photographs were graded as 1 (or "close") when the rostral length represented at least one half of the length of the image (30 cm or more on a 40 x 60 cm projected image), 2 ("medium", rostral length between 10 and 29.9 cm) or 3 ("far", rostral length less than 9.9 cm). These three "distance categories" were used to compare the whale-to-photographer distance distributions between HQ and LQ photographs. This was done to evaluate a possible distance effect in the age predicting capability of the models. The age distribution of whales was also compared between the two samples.

The photographs in the catalog were contributed by 24 organizations and 55 individuals (Hamilton and Martin, 1999) and were taken over many years with different cameras and lenses from a variety of platforms and in different formats (Kraus *et al.,* 1986). For this reason, the sources of error at the moment of photographing the whales varied, and an estimation of this error was practically unfeasible.

Ratios and data analysis

In this paper, the terms "estimate" and "predict" will be used interchangeably. In standard statistical terminology, the word "prediction" involves an inherent uncertainty that is ascribed to each individual whale, as opposed to an average over all whales whose age is "estimated" using the same explanatory variables. The application envisioned here is to estimate (or "predict" according to statistical terminology) the age of whales on the date they were photographed.

Using allometric ratios to estimate the age of whales is particularly convenient because ratios are independent of the distance to (or altitude above) the whales at the moment of photographing them, and because they do not need to rely on a scale or object of known size. The measurements taken on the photographs (Fig. 1) were combined to produce 33 ratios to describe the allometric changes in the head of right whales. Of these ratios, three involved the lip apex; four were designed to measure changes in rostral length; four measured changes in the coaming; 11 combined the length of the rostral islands with rostral length, and 11 were combinations of the measurements involving the curvature of the rostral profile (those involving the zenith).

A study of intra- and interobserver errors was conducted to evaluate the precision of the measurements taken. Measurements of rostral length, coaming length, island length, coming-to-zenith, valley-to-zenith and lip apex-to-isthmus (Fig. 1) were taken 10 times on HQ photographs from each of nine individual whales of different ages by four readers. The coefficients of variation (CV) for the measurements taken by each reader on each trait were calculated to quantify the intraobserver errors. A two-way non-parametric ANOVA was computed using the mean, standard deviation (SD) and CV, respectively, from the 10 repeated measurements for each reader on each trait of each whale. The two factors were reader and whale. A separate analysis was conducted for each of the six traits.

Multiple regression analysis was performed with all photographs and separately within the HQ and the LQ subsamples using the REG Procedure of SAS version 8.2 (SAS Institute Inc., 2001). Early in the analysis it was found that the use of the square root of age as the response variable (instead of age) led to homogeneous variances as required by standard regression. Exploratory regressions were used to fit linear relationships between the square root of age as dependent variable and the 33 ratios as independent variables. Subsets of the best predictors (based on their higher frequency of occurrence in the models with the largest R^2 in the exploratory analysis) were used in further analyses. The REG procedure requires that all independent variables in each observation have a value for the observation to be included in the analysis. Because some measurements were not possible in some photographs, the ratios involving those measurements could not be calculated and therefore, those photographs were excluded from the analyses by the procedure. As a consequence, sample sizes for each group of photographs vary.

To determine if there were sex differences in age prediction, we plotted the true age *versus* predicted age for females and males and calculated the corresponding residuals. The plots were examined, and the means of the residuals for each sex were compared in the HQ photographs. In addition, by using the prediction model for the HQ photographs, we added an indicator variable for gender to explore whether gender increased the prediction capability of the model. Since multiple observations were obtained for the same individual whales, the residuals for each individual were analyzed to determine if there were any consistent "whale effects".

To aid in the validation of the model for the HQ photographs, we randomly divided the photographs into four independent quarters. We used three quarters of the photographs to fit the model and predicted the age of the whales in the remaining quarter of the photographs. We repeated this process four times (one for each quarter withdrawn). We ran a Separate Slopes Analysis (ANCOVA) to test for differences in the relationship between true and predicted age among the four quarters.

To provide an estimate of age prediction error, we obtained the mean and SD for each of the four predictor variables in the HQ sample model. Using all possible combinations of the three values (mean, mean + 1 SD, mean - 1 SD) for the four variables, we fit the model to obtain predicted values of the square root of age with the corresponding 95% confidence intervals (CI) for the predicted values. The values of square root of age were then transformed into true values of age to calculate the CI.

Chi square and Fisher's Exact tests were used to compare the age distribution of whales and the distance distribution in the HQ and LQ photographs. Tests were two-tailed and the level of significance was set at $\alpha = 0.05$.

RESULTS

An average of 6.02 slides (SD = 2.69, range = 2-13) for each of the 30 females and 32 males at different ages were used. The age of whales spanned from 0 (calf year) to 14 yr (mean age span = 8.34 yr, SD = 2.71). The procedure selected 123 HQ slides, 83 LQ slides (50 in group 1, 18 in group 2 and 15 in group 3), and 22 slides of whales with continuous callosity pattern to fit the models.

High quality (HQ) photographs

Using the REG procedure of SAS (SAS Institute Inc. 2001) with the HQ photographs, the mean squared error (MSE) was minimized (or equivalently, the adjusted R^2 was maximized) using a model with four predictor variables involving only six measurements (*i.e.*, adding more variables led to an increase in the MSE). Rostral curvature and dome height showed the greatest variation with age and thus, were the best age predictors. The six measurements were dome height, coaming length, rostral length and the three measurements involving the zenith. The best four predictor variables (Table 2) were *DH/CL, (DH+VZ)/RL, BZ/VZ* and *CZ/VZ*. All had partial p-values <0.001 except for *CZ/VZ,* which was significant at 0.05. We included all four variables in our final model since we used minimum MSE as our criterion for subset selection. The algebraic sign indicates the effect of an increase in the value of the predictor variable on the predicted square root of age. Thus, an increase in *(DH+VZ)/RL* results in a decrease in square root of age. Caution should be excercised in interpreting the coefficient estimates due to correlation among the four predictor variables (multicolinearity).

The R^2 value of 0.747 was quite high suggesting that about three quarters of the variability in the square root of age was explained by the four predictor variables. The equation can be used as a basis for age prediction in right whales. Thus, with the values for the four allometric ratios obtained from a photograph, the predicted square root of age of the whale photographed, $\hat{Y}_P$, is obtained as

$$\hat{Y}_P = 2.201 + 1.125\, DH/CL - 22.841\, (DH+VZ)/RL + 1.396\, BZ/VZ + 0.821\, CZ/VZ$$

Using standard regression techniques, a 95% confidence interval for $\hat{Y}_P$ led to a half width of approximately 1.12 for independent variables in the observed range. Thus, one can be 95% "confident" that the true square root of age of a whale falls within ±1.12 of the value predicted by the model. However, when the response variable (square root of age) is transformed to age, the width of the CI increases with the predicted age of whales. Some representative approximate 95% CIs for some predicted values of age are (lower bound < predicted age < upper bound)

0* < 0.1 < 1.49

0* < 1 < 4.49

0.21 < 2.5 < 7.3

0.77 < 4 < 9.73

2.33 < 7 < 14.18

3.53 < 9 < 16.97

(*constrained to 0 on biological grounds)

The Separate Slopes Analysis indicated that there were no significant differences in the relationship between true and predicted age among the four quarters of the photographs used to validate the model. When the age of the whales in each quarter was estimated using the model fitted with the remaining three quarters, the resulting four slopes of the true *versus* predicted age were close to 1 (range = 0.91 – 1.12) and were not significantly different from each other (ANCOVA, $F = 0.88$, df = 3, $P = 0.45$). This suggests that the model is relatively insensitive to the particular set of photographs used, resulting in increased confidence in its application.

An evaluation was conducted comparing the true age of whales and the predicted age based on our model to determine if there were any systematic problems in our prediction equation. Overall the model fit well. There was less variability and higher accuracy in age prediction among calves and juveniles than among older whales (Fig. 2). For the whales available in our data set, there was a slight tendency for the predicted age to be smaller than the true age for older whales. This is most apparent in Fig. 3, where the "true" number of individuals is larger than the predicted in the three oldest age categories. A partial explanation of this phenomenon involves the use of square root of age for regression and the original units for this figure. The effect is sufficiently small to justify formal steps of bias correction.

Residual plots with whales coded by sex indicated that there were no sex-related patterns. The mean residuals obtained from the plot for HQ photographs in Fig. 2 were not significantly different for females and males (one-way ANOVA, $F = 2.07$, df = 1, $P = 0.15$). When we added an indicator variable for gender in the HQ model, the improvement in the R^2 was negligible (from 0.747 to 0.749). The partial P value of this variable in the model was non-significant ($P = 0.33$). Overall, these results indicate that there are no apparent sex differences in the anatomical traits of the head of right whales that were used for age prediction. Similarly, residuals for individual whales did not show any obvious patterns (*e.g.*, most residuals for each whale having the same algebraic sign), suggesting no whale effects on age prediction.

Low quality (LQ) photographs

The same procedure was followed for the three groups of LQ photographs. Using the minimum MSE criterion, five predictor variables were selected for groups 1 and 3 and two variables for group 2. The models are shown in Table 2. All variables had significant partial p-values except for the intercept in group 1. The R^2 values in the three models (0.795, 0.747 and 0.890) were high. However, it should be noted that the sample sizes were relatively low, especially in groups 2 and 3. Therefore, the predicting capability of models based on LQ photographs of these two groups should be taken with some caution.

When all 83 LQ photographs were used to fit one model with five variables, the resulting R^2 value (0.710) was lower than the individual R^2 values obtained in each of the models that discriminate among the three types of LQ photographs. Similarly, when the four-variable model developed for the HQ photographs was used to predict the age of whales in the three groups of LQ photographs, the R^2 values also decreased (0.68, 0.50 and 0.78 for groups 1, 2 and 3, respectively), making evident the need for different models with photographs of different qualities.

As with the HQ photographs, age prediction with LQ photographs was more accurate for young whales (Fig. 2). The models also underestimated the age of older whales in the LQ sample (Fig. 3).

Whales with continuous callosity pattern

We attempted to develop an age prediction model specific for individuals with continuous callosity pattern. However, some measurements could not be taken on photographs of these whales due to the peculiar anatomical characteristics of this type of pattern. Thus, the number of slides used by the procedure to fit the model was quite low (22). Our chosen model, with four predictor variables, had lower R^2 (0.628) and higher MSE compared to the HQ and LQ photograph models, and the partial p-values of the variables were also higher (Table 2). It is interesting to note that this was the only model that incorporated a ratio with the isthmus-apex measurement.

Age and distance distributions

In order to test for a possible relationship between whale age and photograph quality, we compared the age distribution of whales in the HQ and LQ samples. We divided the age of whales into seven categories: 0, 1, 2, 3, 4-5, 6-8 and 9 yr of age or older. We found no significant differences in the age distribution of whales in the two samples (HQ x = 3.78 yr, SD = 3.34; LQ x = 3.74 yr, SD = 3.46; $\chi^2 = 3.002$, df = 6, $P = 0.81$).

Similarly, we compared the whale-to-photographer distance distribution in both samples to test for a possible distance effect. In this case, we compared the three distance categories (1 = close, 2 = medium, 3 = far, as defined above) among the HQ photographs and the three groups of LQ photographs. We found no significant differences in distance distribution among the samples (mode = 2 in all samples; Fisher's Exact Test, df = 6, P = 0.62). The majority of photographs in both samples (78.9% of HQ and 82.8% of LQ photographs) were taken at medium distance.

To analyze the prediction capability of the model for close and far photographs, we divided the HQ sample into two groups of similar size based on distance, using the median of the rostral length to separate close (= 1) *versus* far (= 2) photographs. We fit the original four-variable model developed with the HQ photographs to each of the two groups. There was virtually no difference in the performance of the model between the two groups (group 1 = close, $R^2 = 0.789$, group 2 = far, $R^2 = 0.798$). This again supports the robustness of the model.

Estimation of measurement precision and reader error

Coefficients of variation (CV) were obtained from 10 repeated measurements of six traits taken by four readers on nine different whales. The CVs for the traits ranged from 0.31% to 4.11% (Table 3). The CVs were lowest for the longest traits (*e.g.*, coaming and rostral lengths) and highest for the shortest traits (*e.g.*, coaming-zenith and valley-zenith distances) among all readers. The measurements of the coaming-to-zenith distance showed the highest variation, with CVs ranging from 0 to 10.9% in one reader.

There were no significant differences among readers in the mean values obtained for measurements for four of the six traits (Table 3). The mean of the coaming-to-zenith distance showed differences among observers, likely due to the small values of this trait (x = 3.96 mm). Although the mean rostral length also showed statistical differences among readers (likely due to the small CV), the absolute magnitude of this difference is minimal (mean values of rostral length for the nine whales measured by the four readers were 137.92, 138.01, 138.09 and 138.24mm).

Not surprisingly, there was a highly significant whale effect in the mean values of the six traits measured (Table 3), indicating the anatomical variability among whales as well as differences in whale-to-photographer distance in the photographs used for the precision estimation test.

DISCUSSION

Age prediction, ratios and models

The models allow for age estimation of North Atlantic right whales by allometric measurements of their heads taken from shipboard lateral photographs. The technique is easy to apply and inexpensive, as it does not require sophisticated and costly equipment to take the photographs and measurements. Because ratios are used, the method does not rely on objects of known size on the same frame of the whales photographed (Whitehead and Payne, 1981) or on the distance (or altitude) from the camera to the animals (Best and Rüther, 1992; Angliss *et al.,* 1995; Perryman and Lynn, 2002) to obtain the values for the age predictor variables. Like other techniques developed since the 1970's, this method is non-invasive and does not require disturbing or killing the whales to obtain vital information such as age-related population parameters.

A total of 10 measurements were taken on the whales' heads (Fig. 1). However, the HQ and the LQ group 1 models used only six measurements, and the remaining models used only five. Once the quality of a photograph is determined, a reader needs only to take the necessary measurements for the specific model to predict the age of the whale on that photograph, which simplifies the use of the technique.

The models work particularly well with younger individuals, while age prediction for whales older than 8 yr has wider confidence intervals. Direct application of the models results in a slight underestimation of the age of older whales. Few whales were predicted to be 10 yr old or older, and none was predicted to be 12 yr or more (see Fig. 3). The flattening of the rostrum as the whales grow older (as shown in Fig. 1), with the resulting loss of age-related anatomical changes in this trait, may hinder age prediction for older whales. The mean age at first parturition is estimated at 11 yr for this population (Kraus, 2002). Thus, the method would work best for young whales, allowing for more accurate age prediction of individuals who have not yet reached puberty.

Best and Rüther (1992) suggested that systematic technical biases in aerial photographs of right whales tended to underestimate the size (and therefore the age) of whales. Whitehead and Payne (1981) obtained greater confidence limits for the mean growth curve of southern right whales at older than at younger ages. The marked curvature of the rostrum or, similarly, the notable height of the dome, in calves and young animals (Hamilton and Martin, 1999) appears to be the best anatomical indication of youth in the head of right whales. This is suggested by the fact that rostral curvature (as indicated by ratios including the zenith measurements) and dome height consistently appeared as predictor variables in all of the models, with the exception of the dome height in the continuous callosity pattern model.

The majority of photographs used were taken in the summer (Jul-Sep), when most calves of the year were between 7 and 9 mo old. If more lateral photographs of the same whales were taken throughout their first few months of life (when changes in the rostral curvature are more prominent), they could be used to develop a more accurate age prediction model for young right whales. Also, in our sample, all whales that were younger than 1 yr were classified as being '0 yr old' (the age of whales was rounded to the nearest year). The models predicted these whales to be between 0.25 and 0.37 yr old on average, which is closer than zero to their true age at the time of being photographed. Others have attempted to describe growth in right whale calves by using the ratio of rostral length to body length, but they found no significant positive allometry in this ratio during the first year of life (Whitehead and Payne, 1981; Best and Rüther, 1992). Given that the aim of the present study is to estimate the age of whales to the nearest year, we think that our technique is appropriate for assigning a whale to the calf year when other indications (*e.g.*, its repeated association with an adult female, or time of the year) are not available (*cf* Clapham *et al.,* 1999).

The rostral length expressed as a percentage of total body length has been used to predict age in southern right whales older than 1 yr (Whitehead and Payne, 1981; Best and Rüther, 1992). Rostral length appeared in all of our models, demonstrating its importance to predict age when combined with other dimensions of the whales' heads. Contrary to the HQ and LQ models, where rostral length appeared in combination with the rostral curvature and the dome height, the rostral length appeared in combination with the isthmus-to-apex distance in the model for whales with continuous callosity pattern.

The model for whales with continuous callosity pattern was the only one to include the isthmus-to-apex distance, and was the only model where the dome height was absent from all ratios. The continuous callosity tissue along the midline of the rostrum of these whales may increase the relative height of the isthmus and distort the typical curvature of the rostrum from a lateral view. One result of this is an increase in the isthmus-to-apex distance compared to whales with discontinous callosity pattern, for which this measurement was not a predictor of age. The increase in the isthmus' height may decrease the measurement of the dome height, potentially removing the positive allometry with age that dome height showed in whales with discontinuous callosity patterns. Also, because the bonnet and the coaming are one indistinct callosity with no discrete endpoints separating them, the location of the margins of 'bonnet' and 'coaming' to take the zenith measurements in whales with continuous pattern is approximate and should be taken with caution. Difficulties with the photo-identification of whales with continuous patterns have been described by Kraus *et al.* (1986). It is important to note that the sample size used to fit the model for whales with continuous callosity pattern was quite small, and these effects could disappear with a larger sample.

Right whale callosities enlarge as the head grows (Payne *et al.,* 1983). Islands may be rounder in calves than in adult whales and may elongate as the rostrum grows. Thus, elongated islands could be an indication of older age. No ratios of the island length to rostral length appeared in the models. The width of the islands cannot be measured on lateral photographs. If dorsal (aerial) photographs were used, a ratio of island length to island width as an indication of the islands' 'shape' could be calculated. This shape could potentially show positive allometry with age (longer in older whales) and be used as an age predictor. Also, the absence of island ratios in the models presented here may be due to the lack of rostral islands in whales with continuous callosity pattern (Kraus *et al.,* 1986), and to some whales having few islands that may be on the opposite side of the head as viewed on the photograph used for a particular year. As a result, these photographs had no values for the island ratios and were excluded by the regression procedure for model fitting.

With the exception of the model for whales with continuous callosity pattern, all models had R^2 values higher than 0.74. When the HQ model was applied to LQ photographs, the resulting R^2 decreased. This suggests that different models should be used with photographs of different qualities. Although the forms of the allometric ratios varied among models, a few traits (most notably, rostral curvature and dome height) appeared to be the most consistent age predictors among models. Similarly, when the appropriate allometric relationships are chosen, total length (and therefore, age class prediction) in fin whales can be reliably estimated when direct measurements are not possible (Ratnaswamy and Winn, 1993). While the selection of predictor variables based on maximum R^2 and minimum MSE can be statistically suitable, no model can be positively identified as the 'true' predictor of age.

Sources of error

A variety of errors reduced the accuracy of age prediction, and therefore increased the width of the confidence intervals of the predicted age of whales. As pointed out by Angliss *et al.* (1995), biases in photogrammetry of cetaceans fall into two main categories. A combination of errors associated with photograph quality and reader error (technical biases) and the inherent inter-individual differences among whales (biological bias) increased the variability in age prediction in our study.

The best photographs were those taken at medium distance with the whale horizontal at the surface and perpendicular to the photographer, and showing one side of the entire rostrum back to the post-blowhole callosities. Deviations from this ideal affected the measurements in different ways (*cf* Gordon, 1990). Some of the most common problems related to photograph quality were: some endpoints were not visible (*e.g.*, the post-blowhole callosities were not in the frame or the front tip of the bonnet was underwater); the whale was tilted on one side and the angle did not allow the viewer to visualize the rostrum's profile from a lateral perspective (especially important to determine the zenith); the whale was leaving or approaching the photographer, thus affecting longitudinal measurements; the image was backlighted and the whale was too dark to find the apex; the lower lips were not in the 'closed' position, thus affecting the measurements involving the apex, *etc*.

The determination of the apex had a considerable variability given that the highest point of the lower lip is usually not a discrete 'point' but a wider portion of the lip, *i.e.*, the apex could be placed forward or backward along the lip according to the reader's criterion. This variability in apex placement also affected the position of the isthmus. The isthmus-to-apex distance was included only in the model for whales with continuous callosity pattern.

In some cases, the position of endpoints such as the rear margin of the bonnet and the front margin of the coaming was difficult to determine due to cyamid infestation on the head (Kraus *et al.,* 1986; Rowntree, 1996). This tended to erroneously enlarge the apparent dimensions of callosities. In many close photographs the texture and color allowed the reader to discern between callosity tissue and cyamids, and the endpoints of the callosities used for the measurements were their true endpoints. In photographs taken from afar, the endpoints were considered to be in the midpoint of the cyamid 'ring' that surrounds the callosities (*i.e.*, the midpoint between what appeared to be callosity tissue and the external margin of the cyamid ring). In these cases, the final determination of the callosity outline was made using the composite drawing of each whale's identifying features in the catalog (Hamilton and Martin, 1999).

The precision of the repeated measurements made by the readers in terms of CVs (Table 3) is consistent with the errors reported in similar studies that used photographs of baleen whales to estimate their size and age, where CVs ranged between 0.6 and 4.6% (Cubbage and Calambokidis, 1987; Best and Rüther, 1992; Angliss *et al.,* 1995; Spitz *et al.,* 2000). Readers could also be trained beforehand to improve measurement precision.

A longer whale-to-photographer distance did not necessarily imply that the quality of photographs was lower than at closer distance. In fact, some photographs taken at longer distances were better than those where the whales were too close, because in the latter a higher angle significantly affected some measurements (*cf* Gordon, 1990). However, when whales were too far, the distances measured between points were small, and consequently the relative error in those measurements increased. As described by Best and Rüther (1992), Table 3 shows an inverse relationship between the size of the measurements and their corresponding CVs, indicating that the error made in finding the endpoints is stable and independent of the magnitude of the measurement involved.

One way to counter this problem, especially when measuring small traits (*e.g.*, island length, zenith measurements), is to increase the distance between the projector and the screen. This increases the size of the projected image and should reduce the error. Nevertheless, the benefit gained with a larger image size may not be worth the loss of image definition, which may in turn reduce precision. Measurement quality could be enhanced by digitising the photographs and using specialized software to take the measurements (*cf* Angliss *et al.,* 1995; Perryman and Lynn, 2002). The use of three-dimensional imagery also allows for higher precision, but the added cost and/or logistical complications may outweigh the minor improvement (Cubbage and Calambokidis, 1987; Angliss *et al.,* 1995; Dawson *et al.,* 1995).

Medium distance photographs (as defined in methods) should be preferred over 'too far' or 'too close', unless the close photographs were taken horizontally to the whale (*i.e.*, close to the water surface) to reduce the angle effect. To reduce the error due to distance when measuring sperm whales, Dawson *et al.* (1995) discarded stereo photographs taken from a boat at distances greater than 60 m for their analysis. Close photographs of right whales have the added benefit of allowing for better discrimination between callosity and cyamids, thus reducing this source of error when finding the endpoints of the traits being measured.

Although we did not formally quantify it, the largest contribution to the width of the confidence intervals of predicted age is likely to be the inter-whale variation in the traits measured. Whitehead and Payne (1981) found that the best measurements of body length and rostral length were only slightly better correlated than the worst measurements, concluding that deviations from the regression line were mainly due to variations between individuals and not to measurement errors.

When several photographs of the same whale are available for different years, multiple sets of measurements and multiple age predictions for the same individual can be obtained. This could increase the accuracy of age prediction for a given whale.

Conclusion

We have presented a benign, simple and inexpensive photogrammetric technique to predict the age of N Atlantic right whales based on shipboard photographs. The compatibility of shipboard and aerial photographs of N Atlantic right whales used for matching known individuals has been described by Kraus *et al.* (1986). A combination of shipboard and aerial photographs, when available for the same individuals, could lead to the development of a similar method to predict the age of whales from aerial photographs. This method could be applied also to southern right whales in the southern hemisphere, where research is based mainly on aerial photographs (Whitehead and Payne, 1981; Best and Rüther, 1992).

ACKNOWLEDGEMENTS

We wish to thank Peter Crump for statistical assistance; William Feeny for help with the illustration; Darcy Geddes, Tina Whitney and Sofia Zahed for their participation in the estimation of measurement precision; the right whale research team at the New England Aquarium, Boston; Kim Marshall-Tilas and Tom Tilas, Whale Conservation Institute / Ocean Alliance, Massachusetts; Roxana Sctheinbarg and Diego Taboada, Instituto de Conservación de Ballenas, Buenos Aires; Davis Travel Award and Vilas Travel Fellowship, University of Wisconsin, Madison, for financial support. M. Sironi was supported by a Fulbright Fellowship, Argentina, and by grants from Charles T. Snowdon while preparing this manuscript.

LITERATURE CITED

Angliss, R.P., Rugh, D.J., Withrow, D.E. and Hobbs, R.C. 1995. Evaluations of aerial photogrammetric length measurements of the Bering-Chukchi-Beaufort Seas stock of bowhead whales (*Balaena mysticetus*). *Rep. int. Whal. Commn*. 45:313-324.

Best, P.B., and Rüther, H. 1992. Aerial photogrammetry of southern right whales, *Eubalaena australis*. *J. Zool. (London)* 228:595-614.

Clapham, P.J., Wetmore, S.E., Smith, T.D. and Mead, J.G. 1999. Length at birth and at independence in humpback whales. *J. Cetacean Res. Manage.* 1:141-146.

Cubbage, J.C., and Calambokidis, J. 1987. Size-class segregation of bowhead whales discerned through aerial stereophotogrammetry. *Mar. Mamm. Sci.* 3:179-185.

Dawson, S.M., Chessum, C.J., Hunt, P.J. and Sloten, E. 1995. An inexpensive, stereophotographic technique to measure sperm whales from small boats. *Rep. int. Whal. Commn*. 45:431-436.

Gordon, J.C.D. 1990. A simple photographic technique for measuring the length of whales from boats at sea. *Rep. int. Whal. Commn*. 40:581-588.

Hamilton, P.K., and Martin, S.M. 1999. A catalog of identified right whales from the North Atlantic: 1935-1997. New England Aquarium, Boston, Massachusetts. 27pp. + 382 plates. [Available from the authors].

Kraus, S.D. 2002. Birth, Death and Taxis: North Atlantic Right Whales in the Twenty-First Century. Doctoral Dissertation, University of New Hampshire. 162pp. [Available from the author].

Kraus, S.D., Moore, K.E., Prince, C.A., Crone, M.J., Watkins, W.A., Winn, H.E., and Prescott, J.H. 1986. The use of photographs to identify individual North Atlantic right whales (*Eubalaena glacialis*). *Rep. int. Whal. Commn*. (special issue) 10:145-151.

Payne, R., Brazier, O., Dorsey, E.M., Perkins, J.S., Rowntree, V.J. and Titus, A. 1983. External features in southern right whales, *Eubalaena australis*, and their use in identifying individuals. pp 371-445 *In*: R. Payne (ed.) *Communication and behavior of whales*. AAAS Selected Symposium 76. Westview Press, Inc., Boulder, CO. 643pp.

Payne, R., and Rowntree, V.J. 1984. *Southern Right Whales: A Photographic Catalog of Individual Whales Seen in the Waters Surrounding Península Valdés, Argentina*. Fundación A. Fortabat, Buenos Aires, and Whale Conservation Institute, Lincoln, MA. 19pp + 55 plates.

Perryman, W.L., and Lynn, M.S. 2002. Evaluation of nutritive condition and reproductive status of migrating gray whales (*Eschrichtius robustus*) based on analysis of photogrammetric data. *J. Cetacean Res. Manage.* 4:155-164.

Rowntree, V.J. 1996. Feeding, distribution, and reproductive behavior of cyamids (Crustacea: Amphipoda) living on humpback and right whales. *Can. J. Zool.* 74:103-109.

Ratnaswamy, M.J., and Winn, H.E. 1993. Photogrammetric estimates of allometry and calf population in fin whales, *Balaenoptera physalus*. *J. Mammal.* 74:323-330.

SAS Institute Inc. 2001. SAS version 8.2. SAS Institute Inc., Cary, NC, USA.

Spitz, S.S., Herman, L.M., and Pack, A.A. 2000. Measuring sizes of humpback whales (*Megaptera novaenglia*) by underwater videogrammetry. *Mar. Mamm. Sci.* 16:664-676.

Whitehead, H., and Payne, R. 1981. New techniques for assessing populations of right whales without killing them. pp. 189-209 *In*: Clark, J.G., J. Goodman and G.A. Soave (eds.) *Mammals in the sea*. Vol. III. FAO Fisheries Service No. 5, Rome.

TABLES AND FIGURES

Table 1: Traits of the head of right whales used for the measurements.

Trait	Description
Dome	the raised portion of the rostrum where the coaming sits
Lip apex	the highest visible point of the lower lip on a lateral view
Isthmus	the narrowest point of the rostrum, determined by the intersection between the middle line along the rostrum and the perpendicular line to the lip apex
Valley	the lowest point of the rostral middle line on a lateral view
Zenith	the imaginary line that connects the highest point of the bonnet with the highest point of the coaming on a lateral view (used to describe the curvature of the rostrum)

Table 2. Age prediction models for high quality (HQ) and low quality (LQ, groups 1, 2 and 3) photographs, and for whales with continuous callosity pattern. Adj: adjusted. MSE: mean squared error; SE: standard error. NS: non significant, P-value > 0.05; *significant, 0.01 < P-value < 0.05; **highly significant, P-value < 0.01. Abbreviations for the measurements used in the variables as in Fig. 1.

Sample	n	R^2 (adj R^2)	MSE	Variable	Estimate	SE	t	P
HQ	123	0.747	0.294	Intercept	2.201	0.366	6.02	**
		(0.738)		DH/CL	1.125	0.211	5.34	**
				(DH+VZ)/RL	-22.841	2.05	-11.14	**
				BZ/VZ	1.396	0.317	4.41	**
				CZ/VZ	0.821	0.374	2.19	*
LQ 1	50	0.795	0.269	Intercept	0.656	0.78	0.84	NS
		(0.772)		DH/CL	0.813	0.368	2.21	*
				(DH+VZ)/RL	-18.263	3.307	-5.52	**
				(BZ+CZ)/VZ	1.796	0.368	4.89	**
				DH/CZ	0.217	0.054	4.04	**
				DH/VZ	-0.218	0.052	-4.16	**
LQ 2	18	0.747	0.291	Intercept	2.221	0.53	4.19	**
		(0.714)		[(DH+VZ)/RL]/ [(BZ+CZ)/VZ]	-30.301	5.153	-5.88	**
				DH/BZ	0.435	0.092	4.71	**
LQ 3	15	0.890	0.246	Intercept	16.66	3.297	5.05	**
		(0.829)		(DH+VZ)/RL	-44.351	7.214	-6.15	**
				CZ/VZ	-10.278	2.827	-3.64	**
				DH/VZ	-2.013	0.668	-3.02	*
				DH/BZ	-0.629	0.15	-4.21	**
				(CZ/VZ)/(BZ/VZ)	5.329	1.529	3.49	**
Continuous	22	0.628	0.420	Intercept	1.844	0.717	2.57	*
		(0.541)		BZ/VZ	1.859	0.524	3.55	**
				CZ/VZ	-0.631	0.479	-1.32	NS
				(VZ+BZ)/CZ	-0.734	0.205	-3.57	**
				IA/RL	7.698	6.885	1.12	NS

Table 3. Average values of mean (in mm), standard deviation (SD) and coefficient of variation (CV, expressed as a percentage) for repeated measuremens for six traits. Significance values of the two-way non-parametric ANOVA with reader and whale as effect. NS: non significant, P-value > 0.05; *significant, 0.01 < P-value < 0.05; **highly significant, P-value < 0.01.

	Values			Reader effect			Whale effect		
Trait	**Mean**	**SD**	**CV**	**Mean**	**SD**	**CV**	**Mean**	**SD**	**CV**
Island length	9.32	0.18	2.13	NS	NS	NS	**	*	**
Coaming length	30.92	0.24	0.90	NS	NS	NS	**	NS	**
Rostral length	138.06	0.39	0.31	*	NS	*	**	NS	**
Apex-isthmus	11.17	0.30	3.29	NS	**	NS	**	*	**
Coaming-zenith	3.96	0.16	4.11	*	NS	NS	**	*	**
Valley-zenith	4.57	0.15	3.61	NS	*	NS	**	*	*

Figure 1. Traits of the head of a right whale and measurements taken on lateral photographs (see Table 1 for descriptions). Traits: B: bonnet; C: coaming; I: isthmus; LA: lip apex; LL: lower lip; PB: post-blowhole callosity; RI: rostral island; Z: zenith. Measurements: BA: bonnet to apex; BZ: rear margin of bonnet to zenith; CA: coaming to apex; CL: coaming length; CZ: front margin of coaming to zenith; DH: dome height; IA: isthmus to apex; IL: island length; RL: rostral length; VZ: valley to zenith. 1987: calf year; 1997: the same whale at age ten.

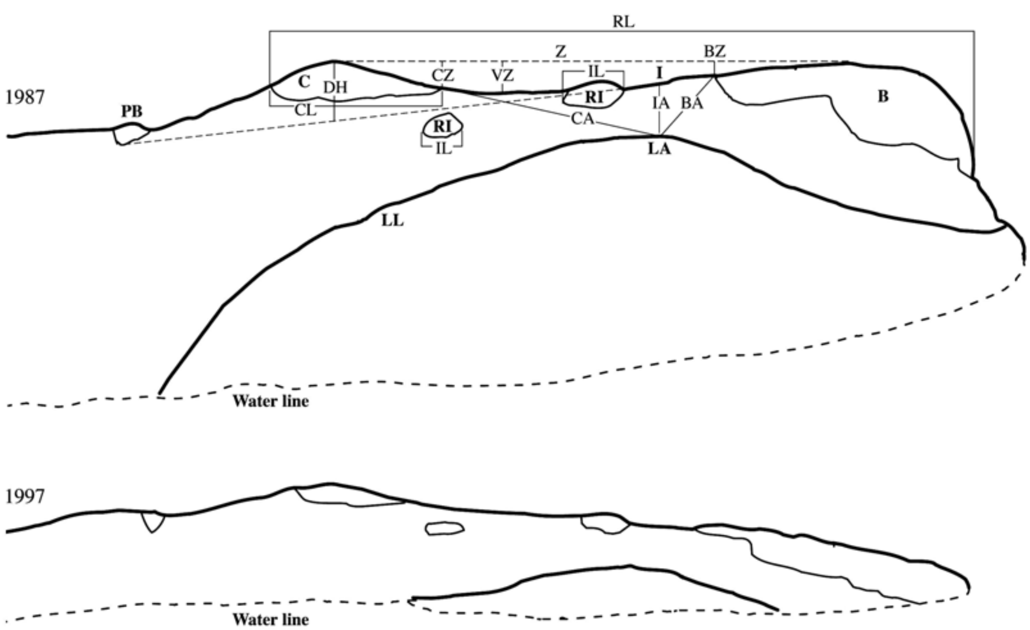
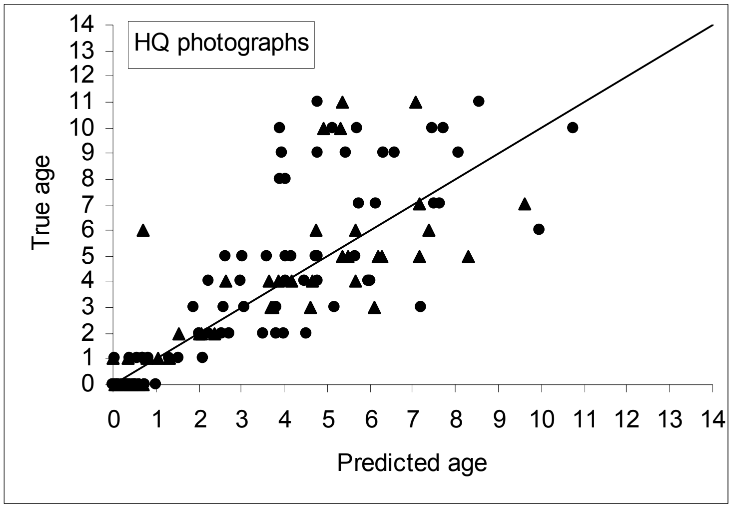


Figure 2. A comparison of true and predicted age (in yr) of females (●) and males (▲) in high quality (HQ) and low quality (LQ) photographs (slope of lines = 1).



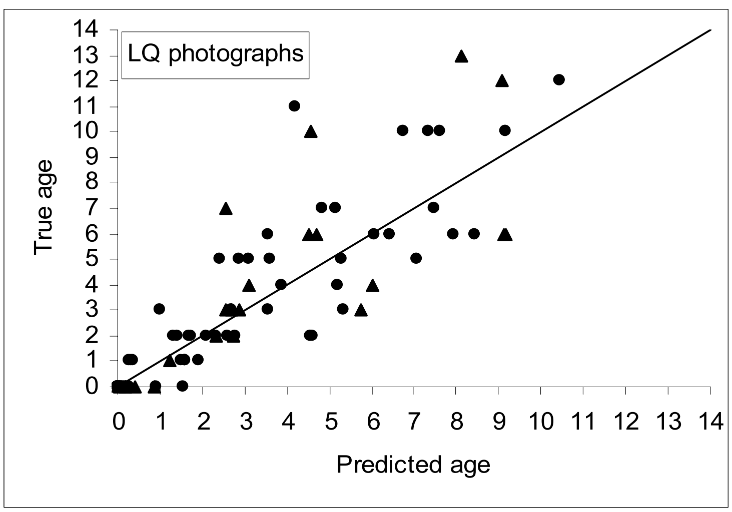


Figure 3. True and predicted age structures (in yr) of the populations of high quality (HQ) and low quality (LQ) photographs.