Migration ranks for bowhead whales (*Balaena mysticetus*) at Barrow in spring

DINARA SADYKOVA[1], TORE SCHWEDER[1,2,*]

[1]*Centre for Ecological and Evolutionary Synthesis, Biological Department, University of Oslo, PO Box 1066, Blindern, 0316 Oslo, Norway*
[2]*Department of Economics, University of Oslo, PO Box 1095, Blindern, 0317 Oslo, Norway*
** email: tore.schweder@econ.uio.no*

ABSTRACT

In a series of aerial photographic surveys of bowhead whales migrating past Barrow in Alaska in the spring, 40 individuals were captured in more than one year. To study individual-specific persistency in migratory pattern, the relative ranks of the captures of these whales among all captures that year are analysed. Controlling for body length and the presence of calves, the correlation of relative ranks in individuals captured multiple times is found not to be significantly different from zero (p-value=0.78).

KEYWORDS: BOWHEAD; MARK-RECAPTURE; MIGRATION; MIXED MODEL; PHOTO-ID;

INTRODUCTION

Bowhead whales in the Bering-Chukchi-Beaufort Seas migrate in the spring north and eastwards past Barrow, Alaska. To what degree does an individual keep its temporal rank in the migration from year to year? This question is interesting from a behavioural point of view. Behaviour is of interest in itself, and is also of concern for abundance estimation and other studies. Schweder et al (2009) utilized results from the present paper when estimating abundance and demographic parameters from aerial photographic surveys of bowhead whales.

Because the spring migration happens during or shortly after the mating season, bowheads have an opportunity to have genetic interchanges across most of the population if there is little temporal stratification in the ranking of whales through the migration. Our question is therefore also of interest when investigating possible structure in the bowhead population in the Bering-Chukchi-Beaufort Seas.

The calendar time of the migratory season varies from year to year (Rugh *et al.*, 2007). The calendar day of capture is therefore not directly useful, and we used the relative rank of a capture within the captures made in the respective year. Provided the surveys were timed similarly relative to the migratory season, the relative ranks are invariant to temporal shifts in the migratory season.

Mothers with calves, and also most large whales without calves, are known to migrate relatively late (Nerini *et al.*, 1984; Rugh, 1990; Angliss *et al.*, 1995). We estimate the effects on relative rank of these covariates, both within all captures and within the recaptures.

We use a mixed effects linear model with normally distributed individual effects, and with fixed effects for length, being associated with a calf, and for differences in years between recapture and capture. The response variable is the logistic transform of the relative rank (rank divided by number of captures in the survey plus one).

METHODS

The data are summarized in Tables 1, 2 and Figures 1, 2. They were obtained through systematic aerial photographic surveys during the spring migration at Barrow, Alaska (Rugh *et al.*, 2007). The length measurements were obtained from the photographic images when possible. When duplicate images have been obtained for a whale, the average value of lengths was considered as a length value for this whale. The mean measured length of the 1,782 captures in the subset where length is recorded is 12.03m and among the 40 recaptured individuals where length is recorded is 13.88m.

To investigate possible persistence over the years in relative rank within individuals, we examined the ranks of the 40 individual whales that were photographically captured in more than one year. The matching protocol is stringent to avoid false positive matches (Rugh *et al.*, 1998, Schweder *et al*. 2009), and we assume that the recorded recaptures are real. There might, however, be many unrecognized recaptures because many bowheads are not marked uniquely enough to be consistently recognized in aerial photographs. The relative ranks at capture and recapture are displayed in Figure 2.

Among the 40 whales, 38 were seen in two years, and two whales were seen in three different years, making 42 recognized between-year recaptures. When length is recorded for both capture and recapture we use its mean for

both captures. This is done to reduce the effect of measurement errors (9 of 28 whales have in fact smaller recorded length at recapture than at capture). When length for one capture or recapture is not recorded, we impute the recorded capture or recapture length of that individual. Any bias introduced by this is small since bowhead whales grow slowly (Angliss *et al.*, 1995; Koski *et al.*, 2006), particularly after sexual maturity.

In addition to body length, an indicator covariate for being associated with a calf, and also the difference in years between recapture and capture are used in a logistic mixed regression of the relative ranks of recaptures which is denoted "time". Large whales are known to be late migrants. The covariate time is introduced to see whether migration tend to be later the older the whale is when controlling for length.

Logistic mixed model

Let years be indexed y and recaptured whales by i. We consider a logistic mixed model with three whale-specific covariates denoted as "calf", "length" as explained above, and "time". The latter is zero at capture and the number of years from capture to recapture. Length is measured in meters, but with mean length for all length-measured individuals subtracted. The response $\mathrm{rank}_{iy} = \frac{\mathrm{r}_{iy}}{\mathrm{n}_y + 1}$ of whale i in year y, where r_{iy} is the sequence number of the capture among the n_y captures that year.

For recaptured individuals the model is:

$$\ln\left(\frac{\mathrm{rank}_{iy}}{1-\mathrm{rank}_{iy}}\right) = \beta_0 + \beta_c \cdot \mathrm{calf}_{iy} + \beta_l \cdot \mathrm{length}_{iy} + \beta_t \cdot \mathrm{time}_{iy} + \xi_i + \varepsilon_{iy}, \qquad (1)$$

where $\beta_0, \beta_c, \beta_l$ and β_t are regression coefficients, ξ_i is a whale-specific random variable that represents the degree of consistency with which whale *i* positions itself in the migration sequence in different years, and ε_{iy} is a residual term. The clusters, i.e. the data referring to individual whales, are assumed independent.

The random individual effects and the residual terms are assumed to be independent and normally distributed with mean zero, and with $\mathrm{Var}(\xi_i) \equiv \sigma^2$ and $\mathrm{Var}(\varepsilon_{iy}) \equiv \theta^2$.

Dependence between any two responses, rank_{iy} and $\mathrm{rank}_{i\hat{y}}$, for the same whale captured in years *y* and $\hat{y}$ respectively can be expressed by the correlation between the logistic transforms

$$\rho \equiv \mathrm{Cor}\left(\ln\left(\frac{\mathrm{rank}_{iy}}{1-\mathrm{rank}_{iy}}\right), \ln\left(\frac{\mathrm{rank}_{i\hat{y}}}{1-\mathrm{rank}_{i\hat{y}}}\right)\right) = \frac{\sigma^2}{\theta^2 + \sigma^2}. \qquad (2)$$

Large individuals, and also mothers with calf, are known to be late migrants (Nerini *et al.*, 1984; Rugh, 1990; Angliss *et al.*, 1995; Koski *et al.*, 2006). How much they delay their migration in terms of relative rank is measured by the logistic regression effects of the covariates calf and length on the relative migration rank for all the 1,782 captures for the subset of the data where length is measured. Here we disregard the fact that 42 of the captures are known recaptures, and use model (1) but with covariate time and random individual effect ξ_i excluded. The index i now runs over all the captures.

A quadratic version of the model is also fitted to the capture-recapture data,

$$\ln\left(\frac{\mathrm{rank}_{iy}}{1-\mathrm{rank}_{iy}}\right) = \beta_0 + \beta_c \cdot \mathrm{calf}_{iy} + \beta_l \cdot \mathrm{length}_{iy} + \beta_t \cdot \mathrm{time}_{iy} + \beta_{ll} \cdot \mathrm{length}_{iy}^2 +$$
$$+\beta_{tt} \cdot \mathrm{time}_{iy}^2 + \beta_{cl} \cdot \mathrm{calf}_{iy} \cdot \mathrm{length}_{iy} + \beta_{ct} \cdot \mathrm{calf}_{iy} \cdot \mathrm{time}_{iy} + \beta_{lt} \cdot \mathrm{length}_{iy} \cdot \mathrm{time}_{iy} + \xi_i + \varepsilon_{iy}. \qquad (3)$$

This model is also used for all the 1,782 captures, but with covariate time (and the linked parameters to that covariate) and random individual effect ξ_i excluded.

To investigate the power of testing for positive correlation, a small simulation study has been carried out. Ranks for all the capture-recaptures are simulated using the logistic mixed model (1), with covariates as observed and with regression coefficients, random individual effects variance and residual variance as estimated. The model was fitted to the simulated ranks and observed data exactly as it was fitted to the observed data. One-sided

likelihood ratio test was performed to calculate p-values for each of 1000 replicates. Finally, GLM on the logistic scale was applied to get a power curve for testing the null hypothesis of $\rho = 0$. Additional repeated simulations were carried out for σ =1.69 (ρ =0.5) and σ =5.1 (ρ =0.9) to get maximum likelihood estimates for the logistic mixed model with simulated ranks. In addition, simulated results were used to estimate a confidence curve and obtain a confidence interval for ρ as in Schweder *et al* (2009).

RESULTS

The model was fitted by way of the computer package AD Model Builder (Otter Research, 2004). Maximum likelihood estimates and standard errors based on the Hessian of the log likelihood are given in Table 3 for the simple logistic regression and the mixed model (1).

The correlation coefficient between ranks (2) is estimated to be 1.9e-12, and is not significantly different from zero (p-value 0.78). Simulation results estimated ρ to be 0.00 with 95% confidence interval (0.00, 0.35). The confidence curve is given in Figure 4 (right panel).

The estimated intercept is higher when only captures with recaptures are considered (Table 3). This reflects that well marked whales tend to be long and thus late migrants.

Being associated with a calf and being long has a significant effect of delaying the migration relative to the other migrants, while the time has no significant effect on the migration rank (Table 3).

Fixing one after another the regression coefficients β_c and β_l to be equal zero for all captures, the likelihood ratio test shows that β_c and β_l are statistically significant (p-value 0.00).

Adding quadratic terms, as in (3), does not improve the fit appreciably for the capture-recapture data. The improvement in log likelihood is only 2.11 units on 5 degrees of freedom. There is thus no evidence for interaction or quadratic effects of covariates. For the capture data of size 1782 the likelihood is improved significantly on 22.67 units, but only the quadratic length term β_{ll} is significant.

DISCUSSION

To illustrate the regression results, take the effect of having a calf. Based on all captures the intercept estimate is 0.17, and the effect of having a calf is estimated to be 2.16. An individual of average length without a calf will thus have predicted relative rank $\exp(0.17)/(1+\exp(0.17)) = 0.54$ while the predicted relative migration rank would be $\exp(2.16+0.17)/(1+\exp(2.16+0.17)) = 0.91$ if it was associated with a calf. Recall that average length (12.03m.) is subtracted from the observed length. The effects on calf and length on relative migration rank are highly statistically significant in the total capture data. In the capture-recapture data the effect of length is strongly significant. The effect of being associated with a calf is only borderline significant here, but recall that there were only 5 cases of capture-recaptures where a calf was associated (Figure 2).

Our results are valid on the assumption that the photo-surveys cover the migration periods in a consistent pattern. If say the first few surveys were performed early in the migration, while the last few surveys were only able to cover the latter part of the migration, the relative ranks would measure different things, and we would have observed an artefact indicating a tendency of the recaptures to appear later in the migration than the captures, even when there is no such pattern in the true ranks. However, except for 1984, 1986, 1987 and 1994, the remaining surveys covered the migration periods consistently (Koski *et al.*, 2006). Since we consider only captures where length of the individual is measured, the surveys in 1984 and 1987 are excluded automatically. Elimination of years 1986 and 1994 produces nearly the same results (β_c, β_l and θ are estimated as 2.18, 0.28 and 1.68 accordingly). Elimination of these four years, when only captures with recaptures are considered, produces similar results for the parameters β_l, β_t, θ and σ (0.96, 0.07, 1.61 and 1.4e-06 accordingly), while the estimate for β_c is different (-1.34), that is no wonder since 4 records with calves were removed from the 5-records database.

Although the capture-recapture data is of small size, the power of detecting migratory patterns is not that bad. The power of testing $H_0 : \rho = 0$ at level 0.05 is about 50% when $\rho = 0.28$. Figures 3 and 4 illustrate the power of testing for positive correlation. Figure 2 and the confidence curve shown on Figure 4 (right) give additional support to our finding of low intra whale correlation in relative rank.

To check the quality of the linear model, it was used as a predictive tool: first the model was applied only to year 1985 and the maximum likelihood estimates obtained from this fit, were applied to the year 1986 to get predicted ranks and residuals. Then the model was fitted to both years 1985 and 1986 to get predicted ranks for year 1989 (years are not consecutive, Table 1), etc. Finally, the model was fitted to all the years 1985-1992 to get predicted ranks for year 1994. The maximum likelihood estimates based on the sequentially cumulated data used to find predicted ranks, are found in Table 4. It is reassuring that these estimates vary little and hardly show any trends. Predicted ranks, residuals and normal-probability plot of the sequential residuals are found in Figure 5.

Except for the well-known systematic effect of large whales, and also cows with calves, migrating late, age is not found to significantly affect the migration rank. The main result is however that individual whales appear not to have any persistency from year to year in their relative rank when passing Barrow in the spring migration. In the limited capture-recapture data the estimated intra whale correlation in relative rank, when re-controlling for covariates, is indeed small and not statistically significant. The power of attesting for positive correlation at significance level 5% is about 50% when $\rho = 0.28$.

ACKNOWLEDGEMENT

The data were kindly provided by David J. Rugh (National Marine Mammal Lab, AFSC, NMFS, NOAA, Seattle, USA) and William R. Koski (LGL Ltd., environmental research associates, King City, Canada).

REFERENCES

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.* 45:313-24.

Koski, W.R., Rugh, D.J., Punt, A.E. and Zeh, J. 2006. An approach to minimize bias in estimation of the length-frequency distribution of bowhead whales (*Balaena mysticetus*) from aerial photogrammetric data. *J. Cetacean Res. Manage.* 8:45-54.

Nerini, M., Braham, H., Marquette, W. and Rugh, D. 1984. Life history of the bowhead whale, *Balaena mysticetus* (Mammalia: Cetacea). *J. Zool.(Lond.)* 204(4):443-468.

Otter Research. 2004. An Introduction to AD MODEL BUILDER *http://otter-rsch.com/admodel.htm*

Rugh, D. 1990. Bowhead whales reidentified through aerial photography near Point Barrow, Alaska. *Rep. int. Whal. Commn.* (special issue) 12:289-294.

Rugh, D.J., Zeh, J.E., Koski, W.R., Baraff, L.S., Miller, G.W., and Shelden, K.E.W. 1998. An improved system for scoring photo quality and whale identifiability in aerial photographs of bowhead whales. *Rep. int. Whal. Commn.* 48:501-512.

Rugh, D., Koski, W.R. and George, J.C. 2009. Interyear re-identification of bowhead whales during their spring migration past Barrow, Alaska, 1984-2004. *J. Cetacean Res. Manage.* (In Press)

Schweder, T., Sadykova, D. Rugh, D. and Koski, W.R (2009) Population estimates from aerial photographic surveys of naturally and variably marked bowhead whales. *J. Agr. Biol. Environ. Stat.* (In Press)

Table 1. Number of captures, number of whales that were re-captured later (First), number of whales captured for the second time (Second), number of captures where length of the individual is measured (Length) and where a calf is associated (Calf).

Year	Captures	First	Second [1]	Length	Calf
1984	21	1	0	0	0
1985	792	18	0	501	4
1986	552	12	7	236	26
1987	365	4	0	0	0
1989	482	3	11	188	22
1990	463	1	8	224	17
1991	426	2	5	243	16
1992	443	1	10	214	1
1994	250	0	1	176	0
Total	3,794	42	42	1,782	86

[1]Two individuals that were captured in three different years, and give two pairs of recaptures each.

Table 2. Bowhead whales resighted during the spring migration past Point Barrow, Alaska. Whale numbers are as defined in the database. Date at capture (Sighing 1), date at recapture (Sighting 2), length at capture (Length 1) and length at recapture (Length 2) are given.

Whale no.	Sighting 1	Length1	Sighting 2	Length2
1058	5/18/1985	13.46	5/13/1986	12.88
1184	5/7/1987	NA	4/23/1992	15.17
1880	5/8/1991	13.59	5/13/1992	NA
1890	5/8/1984	NA	4/23/1992	13.37
1921	5/2/1985	10.15	5/5/1990	10.55
1937	5/11/1985	13.23	4/27/1992	13.36
2024	5/14/1985	10.37	5/12/1986	NA
2037	5/17/1985	14.97	5/29/1986	15.17
2200	5/22/1985	16.31	5/26/1991	16.16
2217	5/23/1985	14.63	5/10/1991	14.49
2246	5/26/1985	13.39	5/6/1989	14.05
2247	5/26/1985	13.38	5/17/1989	NA
2291	5/27/1985	13.50	5/18/1989	NA
2312	5/29/1985	14.59	5/19/1990	13.71
2347	5/31/1985	14.56	5/11/1986	14.67
2371	6/1/1985	15.05	5/26/1992	15.46
2374	6/1/1985	13.88	5/29/1986 [1]	14.29
2384	6/2/1985	12.97	5/15/1989	14.01
2392	6/2/1985	14.45	5/22/1986	13.97
2392	5/22/1986	13.97	5/18/1989 [2]	14.66
2403	6/2/1985	14.34	5/19/1986	13.98
2428	6/6/1985	16.70	5/27/1989	16.01
3963	5/11/1986	9.80	5/14/1992	11.26
4020	5/11/1986	13.33	5/6/1989	13.79
5149	5/9/1992	13.57	5/25/1994	14.45
7946	5/6/1986	12.99	5/3/1989	13.60
8002	5/11/1986	13.44	5/10/1991	14.17
8015	5/11/1986	13.51	6/2/1990 [1]	13.80
8026	5/11/1986	NA	5/16/1992	13.71
8033	5/11/1986	14.60	5/19/1990	14.72
8090	5/14/1986	NA	4/19/1989	12.84
8135	5/22/1986	13.65	4/21/1989	13.05
8142	5/22/1986	13.78	5/19/1990	13.34
8250	5/4/1987	NA	5/11/1990	14.56
8288	5/8/1987	NA	5/25/1991 [1]	16.03

8312	5/18/1987	NA	5/11/1990	13.76
8622	5/19/1986	13.55	5/26/1989 [1]	13.94
8744	4/20/1989	12.61	5/13/1992	13.56
8824	4/25/1989	12.75	5/14/1992	14.57
9304	5/31/1989	NA	5/29/1990	14.18
9304	5/29/1990	NA	5/10/1991 [1, 2]	14.95
10573	5/11/1991	14.68	5/26/1992	15.26

[1]Accompanied by a calf
[2]Third resighting

Table 3. Maximum likelihood estimates and standard errors (in parenthesis) for all captures where length of the individual is measured and for captures with recaptures.

	All captures *n*=1,782	Captures with recaptures *n*=42[1]
β_0	0.17 (0.04)	1.01 (0.24)
β_c	2.16 (0.19)	1.55 (0.79)
β_l	0.26 (0.02)	0.53 (0.15)
β_t		0.04 (0.08)
σ		2.3e-06 (0.01)
θ	1.65 (0.03)	1.69 (0.13)
- log likelihood	1,784.94	86.10

Table 4. Maximum likelihood estimates based on increasing subsets of the data.

Subset	85	85-86	85-89	85-90	85-91	85-92
β_0	0.18	0.12	0.15	0.13	0.11	0.16
β_c	3.31	2.05	2.19	2.38	2.19	2.16
β_l	0.2	0.22	0.25	0.26	0.26	0.27
n	501	737	925	1149	1392	1606

[1] The 2 whales captured in 3 different years were regarded as 4 cases of recaptures.

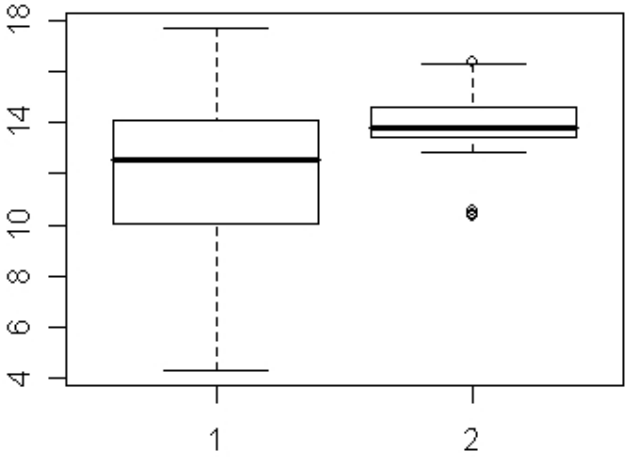

Figure 1. Boxplot of length data for all captures (1) and for captures with recaptures (2).

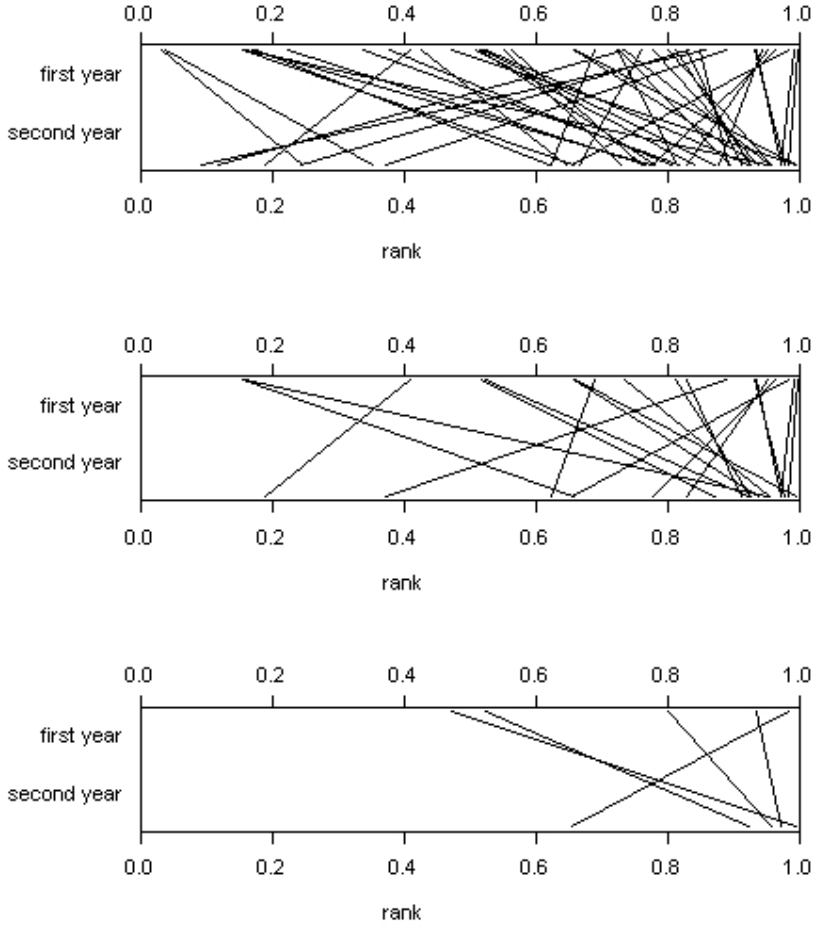

Figure 2. The relative ranks at capture (first year) and recapture (second year) for all the whales (top); for big whales (whales with length more than average 13.88m) (middle); for whales with calves (the whales were seen with calves only in the second year) (bottom).

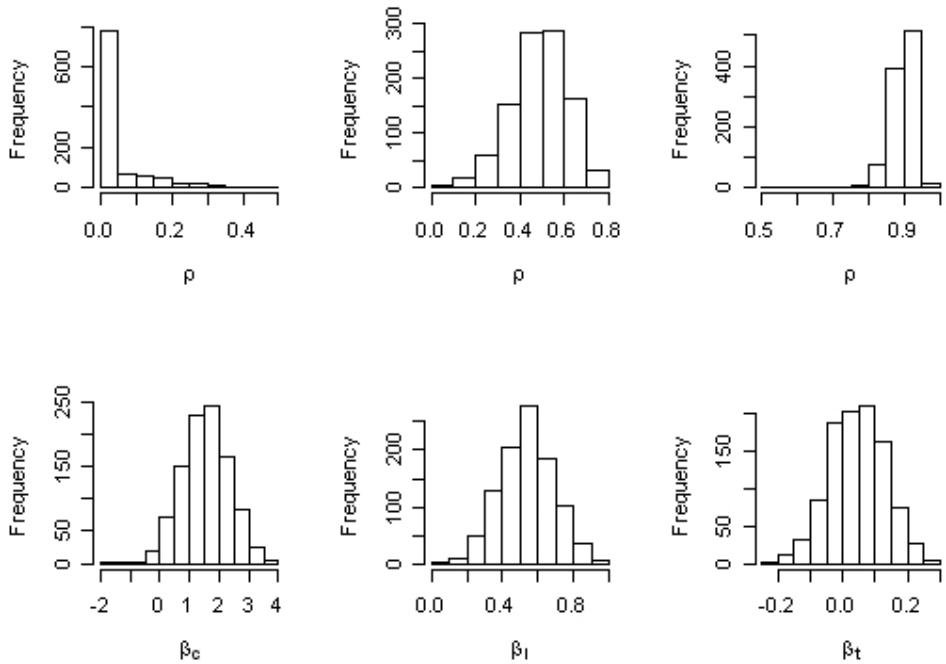

Figure 3. Histograms of 1000 simulated maximum likelihood estimates for the logistic mixed model with simulated ranks for ρ assumed to be: 0.0 (top left), 0.5 (top middle), 0.9 (top right) and for parameters: β_c (true value 1.55, bottom left), β_l (true value 0.53, bottom middle), β_t (true value 0.04, bottom right) for the assumed value of ρ = 0.0.

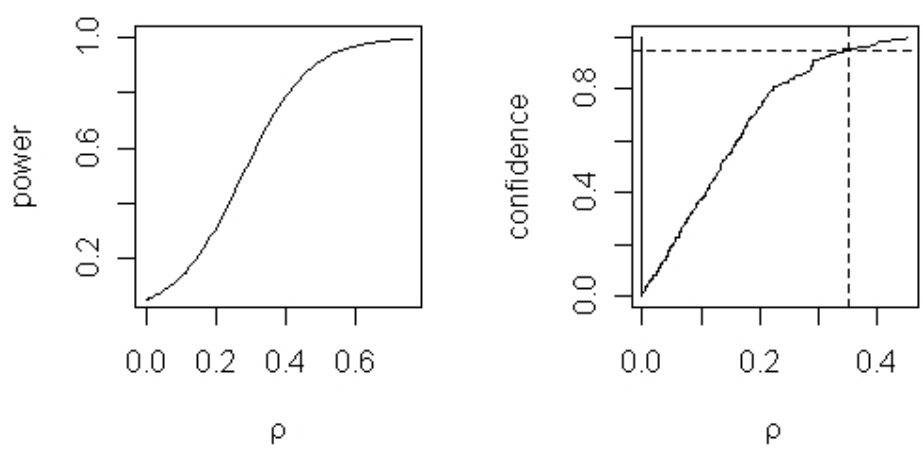

Figure 4. Power curve of testing for positive temporal correlation in relative ranks within individuals at significance level 0.05 (left), and confidence curve for ρ (right). The horizontal dashed line represents the 95% confidence interval (0.00, 0.35). Confidence intervals at other levels of confidence are found as the horizontal line between the two branches of the curve at the chosen level.

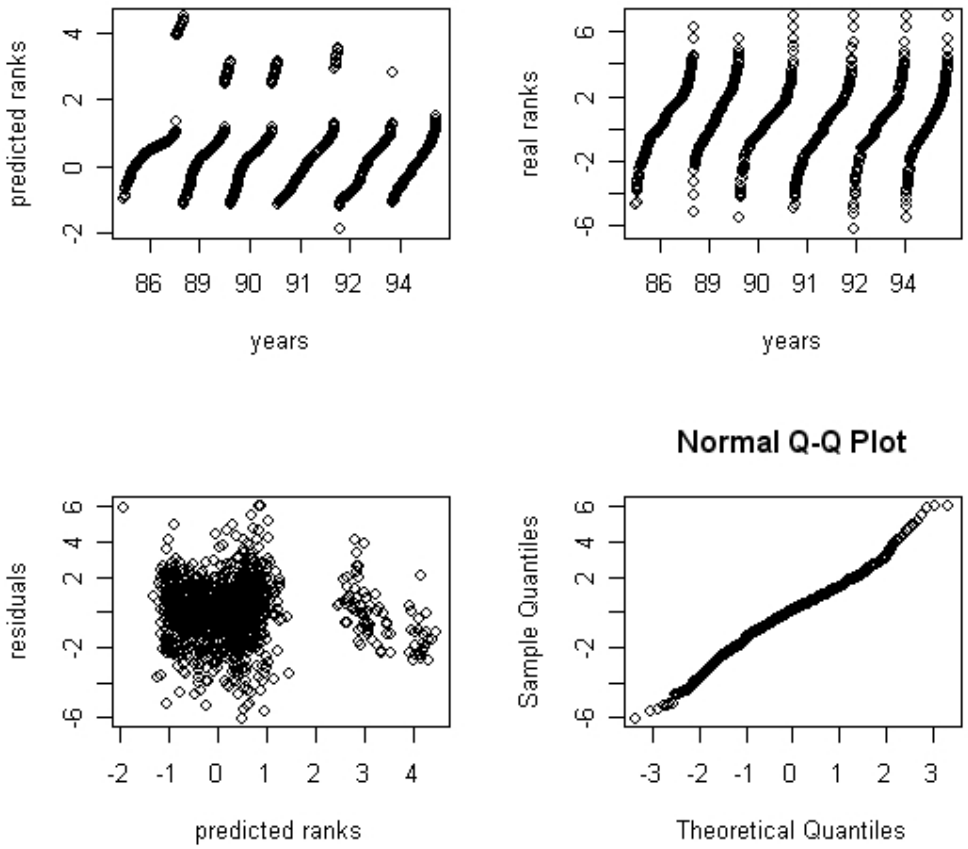

Figure 5. Predicted ranks (sorted) versus years (top left); real ranks (sorted) versus years (top right); predicted ranks versus residuals for all years (bottom left); normal-probability plot of residuals for all years (bottom right).