An analysis of Mute Swan, *Cygnus olor*, breeding data.

P.J. WALTER, P.J. BACON and J.SEARS

Data for Mute Swans breeding in the Oxford Study area during 1977-81 were first analysed by Birkhead et al.*(1983). With the advantage of extra data for 1983-88, their models are reviewed and new features incorporated in the current analysis. The date of lay of the first egg of a clutch is shown to be dependent on the mean winter temperature and the pen's experience. The clutch size is shown to be dependent on the date of lay. The proportion of cygnets fledged from a clutch is shown to be dependent on the date of hatching and both on the pen and cob genotypes.*

Factors affecting the breeding success of Mute Swans *Cygnus olor* were investigated in a preliminary study by Birkhead *et al.* (1983). In their analysis of data for Mute Swans nesting in Oxfordshire, southern England from 1977-81 they examined the effect of a number of predictor variables (such as age, genotype, river type, year and mean winter temperature) on measures of breeding success (laying date, clutch size and number of cygnets fledged). The results suggested a causal chain where; a) female genotype, winter temperature and a pair effect significantly affected the date of laying; b) date of laying and a pair effect significantly affected clutch size and c) clutch size was the only significant factor determining the number of cygnets fledged.

In this paper we re-assess those preliminary findings with the benefit of analysis of a larger data set (1977-88) which enables us to consider extra factors.

Study area and methods

The Oxfordshire Study Area in southern England consists of the River Thames from Lechlade (51° 41'N 1° 41'W) to Goring (51° 31'N 1° 08'W) and three of its tributaries; the Windrush, Cherwell and Thame with adjacent ponds, lakes and gravel pits included (Fig.1.). Data were collected for breeding Mute Swans by the methods described in Bacon (1980) and Birkhead *et al.* (1983). The data for pairs nesting throughout the study area which have been recorded each year from 1977 are given in Table 1.

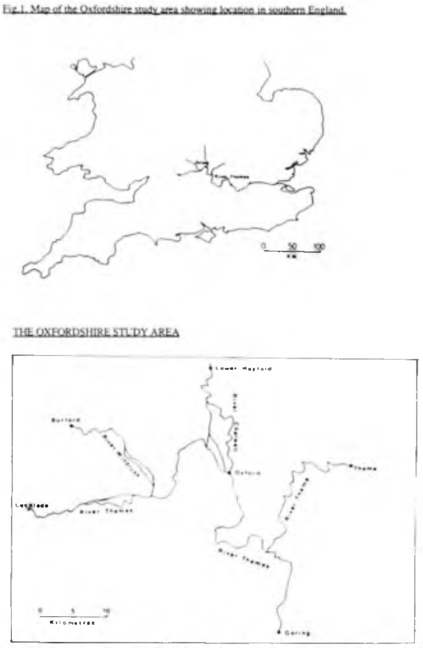

Fig.1. Map of the Oxford study area showing the main rivers.

Problems with analysis

Many records were incomplete for the following reasons: The date of lay, clutch size, number of cygnets fledged and ages of adults are not always known. If a pair loses the whole clutch the adults may not be caught or observed at fledging so the weights are unknown. Although most territory boundaries remained relatively

Supplement No. 1 (1991): 151-156

Table 1. Data recorded for pairs of Mute Swans in the Oxfordshire Study Area, 1977-88.

1. *Causal factors:*
   a) For each bird of the pair, the cob (male) and the pen (female): Identity ring numbers (Darvic and BTO)
   Genotype (Esterase polymorphism, classes SS, SF and FF) Year of Hatch (if unknown estimated assuming age 2 at first breeding attempt)
   Whether it is the first known occasion of breeding in that year.
   Weight (in kg.) in September of that year
   b) Year data:
   Mean winter temperature in °C (overall mean of the means of each of the months December, January, February and March prior to breeding in that year)
   c) Site Data:
   River (codes 1-7)
   Habitat code
   Site

2. *Response variables:*
   d) Breeding data:
   Date of lay (date when first egg was laid - taking 1 January as day 1)
   Clutch size
   Number fledged at approximately 20 weeks old

stable throughout the study and through changes of ownership this is not always so and definitions of site become blurred.

An additional problem was the unbalanced nature of the data. Some pairs which bred for several years had many entries in the data whilst others which bred for a short period only appeared a few times. The Restricted Maximum Likelihood method (REML) was used by Birkhead *et al.* (1983) to overcome these problems of missing data and unbalanced design. The REML method is a modified maximum likelihood method for 'estimating intra-block and inter-block weights in the analysis of incomplete block designs with block sizes not necessarily equal' (Patterson & Thomson 1975). Unfortunately, due to technical reasons, the REML programme failed with the current analysis and less rigorous statistics had to be used. The earlier analysis by Birkhead *et al.*(1983) suggested that the factors which require REML for analysis are of relatively little importance, so the regression analysis used in the current analysis is believed to provide a useful indication.

Synopsis of the findings of Birkhead, Bacon & Walter (1983)

(hereafter referred to as BBW)

The examination of the 1977-81 data led to the following models, (summarised in Fig.2.)

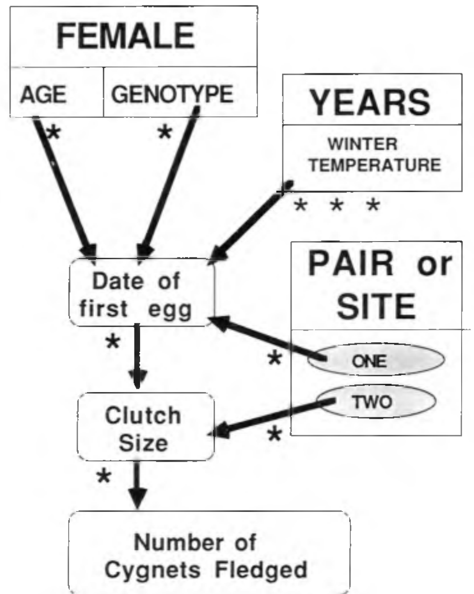

Fig.2. Summary of findings of Birkhead, Bacon & Walter 1983. The response variables (breeding success data) are shown as boxes with round corners; the predictor variables (causal factors) as boxes with square corners. Arrows joining the boxes represent likely cause-effect relationships that were established as significant correlations by REML analysis. The level of significance for each arrow is indicated by the number of * at its base: *** = $P<0.001$; ** = $P<0.01$; * = $P<0.05$. See text for detailed explanation.

A

$$\text{Date of Lay} = 129 - 6.3^{*}\,(\text{Mean Winter temperature}) + \begin{cases} -8.8 & [\text{if SS}] \\ -4.7 & [\text{if SF}] \\ 0 & [\text{if FF}] \end{cases} + p + e$$

p is the swan pair effect and e the error (both assumed to be normally distributed).

The mean winter temperature was highly significant ($P<0.001$).

The pen genotype was only significant at the 5% level which might have been a result of the small data set.

The REML method was necessary to estimate p, and accordingly the genotypic effects. However the estimated magnitude of p was so small that one can be confident that less rigorous techniques such as regression analysis would not have been seriously misleading.

B

$$\text{Clutch size} = 14 - 0.07^{*}(\text{Date of Lay}) + p' + e'$$

p' is the pair effect and e' the error (both assumed to be normally distributed).

The date of lay was highly significant

(P<0.001). The same comment can be made about p' as about p in the previous equation.

C

Number of cygnets fledged
= -0.08 + 0.54*(Clutch size) + e"

e" is the error (assumed to be normally distributed).

The clutch size was highly significant (P<0.001). The constant could be omitted without serious loss of accuracy.

Other factors investigated by BBW

The BBW analysis investigated whether habitat differences had any effect on breeding success over and above the effects of the individual birds breeding on the territory. The REML technique was critical to this, permitting the effects and the variance of individual pairs to be estimated separately from other effects. No significant contribution of habitat type was found.

New considerations in the re-analysis.

The longer run of data allows us to look at breeding experience. Minimum breeding experience is defined as the number of years from when the bird first entered the study to the appropriate year unless it was otherwise known. The pair's experience has been defined as the sum of the experience of both individuals. (This could err on the high side because some pairs do not breed every year.)

We have reconsidered the stage defined by the relationship between the number in the clutch and the number fledged and propose to try to investigate this further. For biological reasons it would seem sensible to investigate the number of eggs in the clutch just prior to hatching, the number of eggs which actually hatch and the date of hatching. In the event there was only time to try the latter. For further biological reasons for considering these points see Bacon & Beekman (1991).

In the BBW analysis only the mean winter temperature was available: we have taken the opportunity to estimate the significance of some other temperature variables.

Table 2. The date of lay predicted from the mean winter temperature and the pen's experience.

	Mean Winter Temperature °C						
	1.5	2.5	3.5	4.5	5.5	6.5	
Pen's Experience							
1	108	103	99	94	90	85	
2	107	103	98	94	89	85	
3	106	102	97	93	88	84	
4	105	101	96	92	88	83	
5	105	100	96	91	87	82	Date of
6	104	99	95	90	86	81	Lay
7	103	99	94	90	85	81	
8	102	98	93	89	84	80	
9	102	97	93	88	84	79	

Methods of current analysis

It needs to be reiterated that a major difficulty in the analysis of the data is that there are measures of some of the variables which are repeated in an unbalanced way. The variables concerned are either year or the corresponding temperature, the genotype of a swan and any value associated with a swan pair. To illustrate this problem consider a swan of genotype FF which appears only once in the study and one of genotype SS which appears three times. The latter will give three times as much information as the former and may therefore give a bias in favour of the information it is giving. This needs to be considered before interpreting the following results too exactly.

The REML method of analysis was not available because of technical problems with the package. The present analysis is therefore based on regression techniques and is preliminary.

Results

A. Date of Lay
The most satisfactory result was:

Equation 1:
Date of Lay = 115-4.5*(Mean Winter temperature) -0.79*(Pen experience) + e

(t-value for -4.5 is -6.61 ***)
(t-value for 0.79 is -3.43**)
(*** <1%; ** <2%; * <5% significant)

Variance explained 16.1%

Regressions with the mean temperatures for the individual months were tested but the most significant relationship was that using the overall mean. We like to infer from this that it is the body reserves accumulated over the whole winter which contribute to a pen's egg-laying potential. No feature of this equation suggests an early limit to laying other than that imposed by the actual ranges of winter temperature and experience. The pen genotype FF was marginally significant but as there were so few birds of this type (about 10%) the result may not be sound. Cob genotype and experience were found not significant.

Using this model to predict the day on which the first egg is laid from the mean winter temperature and the pen's experience one obtains the results given in Table 2.

B. Clutch Size

For the second stage:

Equation 2:

Clutch size = 13.4 - 0.069*(Date of Lay) + e'

(t-value for -0.069 is -11.28***)

Variance explained 28.8%

The most striking point about this stage is that no other predictor variable showed any sign of significance, not even the square of the date of lay. None of the variables genotype, experience or temperature seemed to be important. However, assuming the largest clutch size is 11 this corresponds to a date of lay of 35 which is unrealistically early, so there must be a more satisfactory model.

C. Cygnets Fledged

For the third stage there appeared to be several possible models. A straight forward prediction can be made as follows:

Equation 3:

Number of Cygnets fledged = 0.50*(Clutch size) + e"

(t-value for 0.50 is 25.1***)

Variance explained 15%.

In order to look at the intermediate stages the hatchdate was estimated by the formula:

Hatchdate = Date of Lay + 2 * (Clutch size - 1) + 35.

Using this relationship a more complex and interesting model can be estimated:

Equation 4:

Number of Cygnets fledged

=1.17 + 0.56* (Clutch size) -0.000091*Hatchdate^2) + { 0 [if Pen SS] / 0.61 [if Pen SF] / 0.012[if Pen FF] } + e"

(t-value for 0.56 is 6.4***)

(t-value for -0.000091 is -2.1*)

(t-value for 0.61 is 2.2*)

(t-value for -0.012 is -0.02)

Variance explained 16.8%

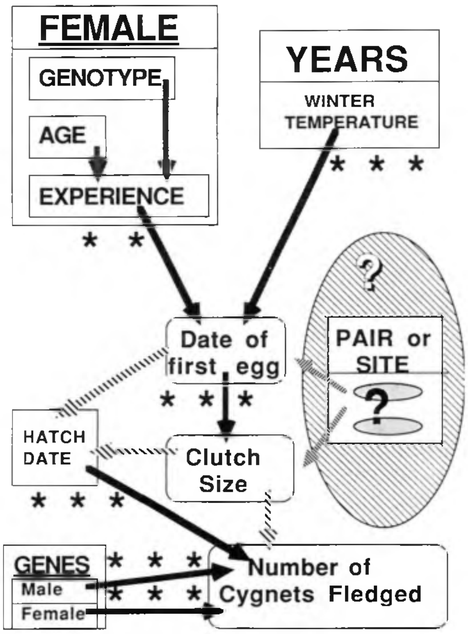

Fig.3. Summary of present analysis. The response variables (breeding success data) are shown as boxes with round corners; the predictor variables (causal factors) as boxes with square corners. The pair/site effect could not be investigated by regression analysis: the variable hatchdate is calculated from date of first egg and clutch size; clutch size also appears in the estimation of fledging success via the logit formula of Equation 5. Arrows joining the boxes represent likely cause-effect relationships that were established as significant correlations by REML analysis. The level of significance for each arrow is indicated by the number of * at its base: *** = $P<0.001$; ** = $P<0.01$; * = $P<0.05$. See text for detailed explanation.

This result needs to be treated with extra caution because a number of items in the data do have undue influence. It is also unsatisfactory because one could find reasonable values which predict zero or negative results. Finally it should be noted that the hatchdate and the clutch size are correlated.

An even more interesting model is a so-called logit model which estimates the probability p of a cygnet reaching the fledging stage starting from the egg in the clutch before incubation:

Equation 5:

$\log_e[p/(1 - p)]$

= 1.20 - 0.000064*(Hatchdate^2) + { 0 [if Pen SS] / 0.43 [if Pen SF] / 0.02 [if Pen FF] } + { 0 [if Cob SS] / -0.24 [if Cob SF] / 0.48 [if Cob FF] } + e

(t-value for -0.000064 is -4.36***)

(t-value for 0.43 is 4.36***)

(t-value for 0.02 is 0.11)

(t-value for -0.24 is -2.63***)

(t-value for 0.48 is 2.3*)

Table 3. The probability of cygnets reaching the fledging stage as predicted by the logit model (equation 5) from a given clutch size and date of lay using cob and pen genotypes SS.

	Number of eggs in the clutch										
	1	2	3	4	5	6	7	8	9	10	11
Date of Lay											
70	0.6	0.6	0.6	0.6	0.5	0.5	0.5	0.5	0.5	0.5	0.5
90	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.4	0.4	0.4	0.4
110	0.5	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.3	0.3	0.3
130	0.4	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.2	0.2	0.2
150	0.3	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.1	0.1
170	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

Table 4. The probabilities of cygnets fledging converted into total number of cygnets fledged as predicted using the logit model (equation 5) from a given clutch size and date of lay using cob and pen genotypes SS.

	Number of eggs in the clutch										
	1	2	3	4	5	6	7	8	9	10	11
Day of Lay											
70	0.6	1.2	1.8	2.4	3.0	3.5	4.1	4.6	5.1	5.6	6.0
90	0.5	1.1	1.6	2.1	2.6	3.0	3.5	3.9	4.3	4.7	5.1
110	0.5	0.9	1.3	1.7	2.1	2.5	2.8	3.2	3.5	3.8	4.0
130	0.4	0.7	1.0	1.4	1.6	1.9	2.2	2.4	2.6	2.8	3.0
150	0.3	0.5	0.8	1.0	1.2	1.3	1.5	1.7	1.8	1.9	2.0
170	0.2	0.4	0.5	0.6	0.8	0.9	1.0	1.1	1.1	1.2	1.3

Table 5. The proportion and number of cygnets fledged for a given date of lay as predicted using a combination of equation 2 and equation 5, for cob and pen genotype SS.

Day of Lay	Clutch Size	Hatch Date	Proportion Fledged	Number Fledged
70	8.6	119	0.57	4.6
90	7.2	137	0.50	3.5
110	5.8	153	0.43	2.1
130	4.5	171	0.34	1.4
150	3.1	189	0.25	0.75
170	1.7	205	0.18	0.18

This model is more satisfactory because it predicts probabilities of fledging which are reasonable over the whole range of predictor variables. However there is much more research to be done because here again some particular observations are exerting undue influence on the final model. Taking all these stages together a new causal chain illustrated by Figure 3 is suggested.

The predictions given by the models are illustrated in Tables 3,4 and 5.

Discussion

Clearly these models suggest that the earlier a pen lays the more successful the breeding is likely to be but they do raise several questions and require much more detailed work. Does the fact that the mean winter temperature over the whole winter gives a more satisfactory model than a single month's temperature imply that the whole winter is required to prepare for breeding? Other measures of weather conditions need to be investigated. What is the nature of the experience gained by the older pens which allows them to lay earlier?

Do some of the differences between the genotypes imply a long term change in the population or is there some so far undetected counterbalancing effect? The BBW analysis ascribed a significant genetic effect on predicting the date of lay to SS pens only. Previous analyses by Bacon (1980) and Walter (1981) found nearly significant genetic effects of cob genotype. The present analysis found a smaller effect of genotype (barely significant) on date of lay, but two highly significant effects, for both cob and pen, on fledging probability. We first draw attention to the fact that the magnitudes of the estimated factors predict a significant advantage (+0.43) to SF pens but a significant disadvantage to SF cobs(-0.24). This is a complex situation, but one which could well lead to a permanent maintainance of a balanced polymorphism.

The absence of the genetic effect of pen genotype on laying-date in the current analysis compared with the BBW results is initially puzzling. However the current results show significant effect of pen experience, a factor BBW were unable to include. The apparent anomaly could be explained by the interaction between genotypic effects and 'experience'.

In this investigation we wish to give special thanks to Mike O'Regan of Research Machines Limited, Oxford for loaning the microcomputer on which these analyses were prepared.

References

Bacon, P. J. 1980. Population genetics of the Mute Swan *Cygnus olor*. D.Phil thesis, University of Oxford.

Bacon, P. J. & Beekman, J. 1991. Models of swan population dynamics and growth - a review. In: J. Sears, & P.J. Bacon, (eds.) *Proc. 3rd Int. Swan Symp.* Oxford, 1989. *Wildfowl* Special Supplement No. 1

Birkhead, M. E., Bacon, P. J. & Walter, P. J. 1983. Factors affecting the breeding success of the Mute Swan *Cygnus olor*. *J. Anim. Ecol.* 52: 727-741

Patterson, H. D. & Thompson, R. 1975. *Proc. 8th Int. Biometric Conference* 197-207. Editura Academici Republicii Socialiste Romonia.

Walter, P. 1981. A statistical study of the breeding success of Mute Swans. M.Sc. thesis, University of Oxford.

P. Walter, 16 Abbott Road, Abingdon, Oxon. UK.
P.J.Bacon, ITE Merlewood Research Station, Grange-over-Sands, Cumbria, UK.
J.Sears, EGI, Dept. of Zoology, University of Oxford, South Parks Rd, Oxford, OX1 3PS, UK.