Review and analysis of studies and documents on the topic of waterfowl and low level flights in Labrador and Northeastern Québec

Prepared for
the Institute for Environmental Monitoring and Research

CONTRACT NUMBER 2423

July 2001

DUVETNOR

Société Duvetnor Ltée
200 rue Hayward, C. P. 305
Rivière-du-Loup, Québec
G5R 3Y9

Summary

We review a set of 14 technical documents pertaining to the presumed impact of low level training flights (LLTF) on waterfowl distribution and abundance. None of the studies we reviewed were flawless and none were truly designed to assess the impact of the flights on the birds. The complexity of the problem is extreme and as LLTF have already taken place throughout the area and may already have influenced waterfowl abundance and distribution, a truly experimental approach to the problem is probably no longer possible. Most of the studies reviewed were summaries of incidental observations on waterfowl in connection with hydroelectric development or in connection with understanding continental demographic trends in selected species such as Black Duck and Canada Geese.

Furthermore, many of these studies had inherent weaknesses (for instance, the use of questionable correction factors) and often major flaws in design (for instance, randomization of transects/plots was not deemed essential in some studies while in others, plot size and shape varied inexplicably). We conclude that none of the studies we reviewed has pertinence in assessing the impact of LLTF on waterfowl distribution and abundance.

We conclude our review with recommendations concerning future studies: designing a comprehensive, experimental protocol to assess the impact of LLTF on waterfowl is not possible in our opinion. We rather recommend that waterfowl habitat be mapped throughout the range and this information combined with general ecological knowledge about the various components of the waterfowl community so as to delineate key waterfowl habitats. In a precautionary approach, these habitats could be excluded from LLTF once an appropriate ranking of the various species has been established in terms of vulnerability. Carefully designed specific studies should complement the overall assessment. Such studies should be critically reviewed before they are launched and should be assessed periodically by expert committees.

Table of Contents

Introduction

The Institute for Environmental Monitoring and Research (IEMR) has mandated Duvetnor to critically review 14 technical documents and appraise their usefulness in evaluating the impact of low level training flights (LLTF) on waterfowl populations. The report is divided in three sections. The first is a detailed review of each document followed by a general review section in which we evaluate the coverage of the various studies about distribution, abundance, and status of the various waterfowl species found in the study area throughout the annual cycle. We also assess the survey methodology, particularly the pertinence of the methods used, the abundance estimates, and the variability of these estimates of waterfowl populations and the potential biases underlying these data sets. A final section based on our critical review presents perspectives on future waterfowl research by looking at gaps in knowledge, setting future priorities from a technical point of view, and examining what is needed in terms of study design and statistical analyses to assess low level flight (LLF) effects on waterfowl.

The documents to be reviewed are listed in Table 1 and were classified according to their content and objectives. We first identified two documents that reviewed the 1994 Environmental Impact Statement pertaining to the military flying activities in Labrador and Québec. For these reports, we simply comment on the validity of the different opinions expressed.

A second series of documents (7) was classified as technical reports presenting results from standard aerial surveys. In our detailed review, we only comment about the relevance of each report in assessing low level flight effects. Criticisms on aerial survey methods are discussed in the general review section.

Lastly, we gathered the rest of the documents (5) in a third category that addresses specific issues about the effects of LLF on waterfowl populations or behavior. These documents were scrutinized using the following framework: working hypotheses, observational methods, sampling design, statistical analyses, and interpretation of the results.

Table 1. List of documents reviewed in this report (in chronological order).

Reviews of Environmental Impact Statement

Hogan, H. A. *1994.* ***Review of Environmental Impact Statement – Military flying activities in Labrador and Québec.*** Prepared for the Innu Nation and Conseil des Atikamek et des Montagnais, 22 pp.

RRCS Ltd. *1994.* ***Impact of military flying activities on wildlife, with emphasis on birds.*** Prepared for the Department of National Defence, Renewable Resources Consulting Services Ltd. (RRCS Ltd.), Sidney, BC, 45 pp.

Aerial Survey Reports

Gillespie, D. I. and S. P. Wetmore. *1974.* ***Waterfowls surveys in Labrador/Ungava, 1970, 1971, 1972.*** Pp. 8–18 in H. Boyd (ed.) Canadian Wildlife Service Studies in Eastern Canada, 1963–73. Canadian Wildlife Service Report Series No. 29.

Goudie, R. I. and W. R. Whitman. *1987.* ***Waterfowl populations in Labrador, 1980–82.*** Pp. 45–63 in A. J. Erskine (ed.) Waterfowl Breeding Population Survey, Atlantic Provinces. Canadian Wildlife Service Occasional Paper No. 60, 82 pp.

S. Fudge and Associates Ltd. *1989.* ***Identification of spring staging areas utilized by waterfowl on the Ungava Peninsula.*** Prepared for the Department of National Defence, 10 pp.

Bateman, M. C. *1993.* ***Canada Goose Breeding Ground Survey – Labrador 1993.*** Canadian Wildlife Service, Atlantic Region, 18 pp.

Bateman, M. C. *1994.* ***Canada Goose Breeding Ground Survey – Labrador 1994.*** Canadian Wildlife Service, Atlantic Region, 15 pp.

Goudie, R. I., D. Lemon and J. Brazil. *1994.* ***Observations of Harlequin Duck, other waterfowl, and raptors in Labrador, 1987–1992****.* Canadian Wildlife Service Technical Report Series No. 207, 13 pp. + tables and figures.

Collins B. *1998.* ***Analysis of 1999 Black Duck Breeding Ground Survey.*** Memorandum, Canadian Wildlife Service, Atlantic Region, 2 pp. + tables, figures and appendix.

Table 1 (continued). List of documents reviewed in this report (in chronological order).

Specific studies

RRCS Ltd. *1993.* ***Goose Bay Waterfowl Avoidance Monitoring Program: 1992 Final Report.*** Prepared for Project Management Office – Goose Bay – Department of National Defence, Renewable Resources Consulting Services Ltd. (RRCS Ltd.), Sidney, BC,

Bateman, M. C. and A. H. Hicks. *1997.* ***Progress Report – Monitoring Waterfowl Breeding Densities.*** Prepared for the Department of National Defence, PMO Goose Bay, Canadian Wildlife Service, Sackville, NB, 7 pp. + figures and tables.

Bateman, M. C., A. H. Hicks and S. M. Bowes. *1999.* ***Waterfowl Behavior in Response to Jet Overflights at Snegamook Lake, Labrador.*** Prepared for Goose Bay Office, National Defence Headquarters, Ottawa, Canadian Wildlife Service, Sackville, NB, 139 pp.

Bateman, M. C. and A. H. Hicks. *1999.* ***Waterfowl Populations in the Low Level Training Area of Labrador – A Data Compilation and Analysis.*** Prepared for Goose Bay Office, National Defence Headquarters, Ottawa, Canadian Wildlife Service, Sackville, NB, 75 pp.

Turner, B. and A. Hicks. *2000.* ***Breeding Population Trends of Waterfowl in the Military Low Level Training Area.*** Prepared for Goose Bay Office, National Defence Headquarters, Ottawa, Canada, Canadian Wildlife Service, St. John’s, NFLD, 42 pp.

Detailed Review of Each Document

Reviews of the Environmental Impact Statement

Review of Environmental Impact Statement – Military flying activities in Labrador and Québec
by H. A. Hogan. *1994.*

Under contract with the Innu Nation, Hogan reviewed aspects related to waterfowl that were designated as Valued Ecosystem Components (VECs) in the revised 1989 EIS released by DND in 1994. The main conclusions were the following:

- *DND has not studied any potential cause and effect relationship between LLF and waterfowl distribution and breeding biology.* **True**, but see Bateman *et al.* (1999) and Turner and Hicks (2000) for such recent studies during part of the annual cycle.

- *Need for an ecosystem approach that would include the life history of waterfowl species, especially for less common species like Harlequin Duck.* **True**, creating an avoidance area around the location of one sighting of a pair or female during one survey is a very short-sighted measure. The birds move around and could be km away a week later depending on the period of the year. Similarly, a site with no observation (whether it was surveyed or not) does not mean that it has no potential and that it will never be used. Habitat characteristics may be more informative to establish avoidance areas.

- *Cumulative impacts are not fully addressed.* **True**

- *Option B (increasing the training area and shifting it southward) seems to have a lower impact on waterfowl.* **Maybe**.

- *The use of density thresholds to establish avoidance levels does not appear a suitable approach.* **True**, because of the annual and seasonal variation in density estimates attributed to extrinsic factors (weather, timing of migration and breeding, population dynamics, imprecision of the surveys, etc.).

These comments and critiques seem justified and sound but we have not reviewed the EIS itself.

Impact of military flying activities on wildlife, with emphasis on birds
by Renewable Resources Consulting Services Ltd.*, 1994.*

Provide responses for public hearings to some reviews and concerns about the EIS. The VEC approach appears suitable. They stated that

- *Time and budget constraints prevent the experimental testing of hypotheses for each VEC* (p. 12). **True**, but should be tried for most important ones.

- *The impact ratings are subjective and not based on strong evidence* (p.16). **True**, the consultants acknowledged this.

- *Option B appears less detrimental for wildlife than A.* **May be true** but we have not seen the full details of the EIS, which makes it difficult to judge the validity of this report.

Aerial Survey Reports

Waterfowl surveys in Labrador/Ungava, 1970, 1971, 1972
by D. I. Gillespie and S. P. Wetmore. *1974.*

Reports results from extensive helicopter and fixed-wing aircraft surveys conducted in May, June, and July 1970–1972 in Labrador, including some in the LLF zone. **Not designed to test the effects of LLF** but instead to look at the potential impacts of the future hydroelectric development of Churchill Falls. Plots and transects were used to estimate pair and brood densities. Defined different habitat types according to their waterfowl capability. No correction factors for visibility were used. The pertinence of this paper for the LLF is limited. Data are too old to be used as control because many changes, some outside the LLF area, have occurred since.

Waterfowl populations in Labrador, 1980–82
by R. I. Goudie and W. R. Whitman. *1987.*

Reports results from extensive surveys in Labrador from 1980 to 1982. **Not designed to test the effects of LLF** but instead, the effects of the Smallwood Reservoir.

Helicopter surveys in June 1980 in the same plots and transect lines as Gillespie and Wetmore (1974) plus ground surveys on several plots. Also, transects were done at 30 km intervals across Labrador using fixed winged aircraft, coupled with ground surveys on 15 plots during the same time. Surveys in 1981 were done in a subset of plots. Mentioned ground surveys conducted in 1982 in the Groswater Bay region but did not present the data.

Comparisons with data from the 70s reflected continental changes: increasing Canada Geese and decreasing Black Ducks.

Air and ground survey comparisons were used to provide correction factors. However, they were conducted in a single ecoregion and in different years, which is not reliable. Total estimated populations of waterfowl in Labrador were calculated using these correction factors and thus may be strongly biased. Fall flight forecast was also calculated from these estimates. The expansion factor used to predict fall population is not documented. These criticisms are worthy of mention because some of the documents reviewed used these estimates as a reference.

Identification of Spring Staging Areas Utilized by Waterfowl on the Ungava Peninsula
by S. Fudge and Associates Limited. *1989.*

Reports results from surveys conducted from fixed-wing aircraft during the spring staging period along transects (2880 km) in May 1989. **Not designed to test the effects of LLF.** Survey routes were selected on the basis of the 1988 survey data from the literature (not cited) and after consultation with residents.

Marked differences were observed between years. They identified a series of areas with concentrations of waterfowl, notably the west end of Snegamook Lake. The stated objective of "identifying locations and timing" is somewhat misleading. No information is provided to determine if routes were established to identify other potential ice-free areas or if the authors considered whether they had covered most potential areas. The effect of timing is not really evaluated since only one transect flight seems to have been done on each individual route.

Canada Goose Breeding Ground Survey – Labrador 1993
by M. C. Bateman. *1993.*

A transect survey for breeding Canada Geese (Atlantic Flyway migrant Canada Geese) was conducted in 1993 using a fixed-wing aircraft (6678 km of transect). Visibility bias established in two different years using plots in the "same" region comparing fixed-wing and helicopter data. Author assumed that helicopter surveys provide unbiased results: an unacceptable conclusion. It is known that waterfowl (especially Canada Geese) may escape at the approach of a helicopter, thus leading to underestimates of true density. The correction factor (2x) is thus questionable.

Using a subsample of 15 transects, the author found no difference in pair numbers compared to 1980 but fewer non-breeders. No mention is made of how standard errors presented in tables were determined. Power tests are lacking, which is especially crucial considering small sample size. Higher density of pairs observed in the LLF zone but no statistical comparisons. **The survey was not specifically designed to test the effects of LLF.**

Canada Goose Breeding Ground Survey – Labrador 1994

by M. C. Bateman. *1994.*

Same survey method as in 1993 with increased length of transects in Quebec (total=7714 km). No difference was found in the number of indicated breeding pairs compared to 1993, but more non-breeders were found. Using a subsample of transects (n=15), comparisons were made with data from 1980. Number of pairs was lower in 1994 vs 1980 and equal to 1993 whereas the total number of geese was greater in 1994 than in 1993, but lower than in 1980. Again, it is not mentioned how standard errors presented in tables were determined. Power tests are again lacking which is especially important with such a small sample size. The author mentions the closing of hunting in Quebec and Eastern Ontario in 1994 but this occurred AFTER the survey. No differences were found for the number of breeding pairs and the total number of geese between areas within and adjacent to the LLF zone but no statistical comparisons. **Again, this survey was not specifically designed to test the effects of LLF.**

Observations of Harlequin Duck, other waterfowl, and raptors in Labrador, 1987–1992

by R. I. Goudie, D. Lemon and J. Brazil. *1994.*

The authors summarized helicopter surveys, including those of Ledrew-Fudge & Associates consultant conducted in June 1987 and July 1991 in or near the LLF zone. Few pairs of Harlequin were noted throughout the area. **Not designed to test the effects of LLF.** No random selection of river segments. No correction factors applied to the data. A density of 8.4 pairs/100 km of river (not /100 km^2 as indicated in the abstract) was established and considered low relative to other reported populations.

Analysis of 1999 Black Duck Breeding Ground Survey

by B. Collins. *1998.*

This CWS internal report presents results of the 1999 helicopter surveys conducted as part of the Black Duck Joint Venture. **These surveys were not specifically designed to test the effects of LLF.** The location of the plots and the strata are not presented in this report and could not be related to the LLF areas, although this information exists. More detailed analyses could be done, using plots in the LLF areas if sample size is adequate. On the other hand, annual trends established for a large area encompassing the LLF area could be used as a control to compare results of surveys within the LLF area. This is especially important for species that could have been affected by changes in hunting regulations in recent years (e.g., Black Ducks and Canada Geese). The statistical method described in Appendix A to measure annual trends could be used to analyze data collected as part of a long term monitoring program of waterfowl in the LLF area. This is based on a trustable procedure first developed by Link and Sauer (1994) for establishing annual trends using the Breeding Bird Survey (BBS) data.

Specific Studies

Goose Bay Waterfowl Avoidance Monitoring Program: 1992 Final Report
by Renewable Resources Consulting Services Ltd. *1993.*

Stated objectives

This document reports the results of a breeding bird survey conducted in 1991 and 1992 to locate concentrations of waterfowl to establish zones of avoidance for LLF. **Does not test the effects of LLF.**

Observational methods

Surveys conducted with fixed-wing aircraft at the beginning of June in 21 (1991) and 80 (1992) plots varying in size and shape, which is odd from a sampling point of view. Larger plots were partially covered (40–80%) and no corrections or weighting factors seem to have been applied. There is also a risk that densities were correlated with the area sampled, but this was not verified.

Sampling design

Plot selection was based on several habitat characteristics recognized from maps. It is not mentioned whether plots were randomly selected or simply chosen from an exhaustive list of potential areas identified on the maps. If avoidance criteria were applied to a subset of potential habitats, it does not fulfill its primary purpose of maximizing protection of waterfowl.

Statistical treatment and data analysis

Mann-Whitney U test was used to compare waterfowl densities among plots and two regions, low level training area (LLTA) 1 and 2.

Presentation and interpretation of the results

The presence of non-breeding Canada Geese (1- and 2-year-old birds) could explain the difference between estimates of the total number of birds (0.19/km^2) vs number of birds recorded as indicated breeding pairs (0.16 birds in pairs/km^2). However, numbers of non-breeders were low (0.03 birds/km^2) showing that this segment of the population was not very abundant in the surveyed plots. For dabblers, the difference should be even smaller since most birds usually breed at 1 year old but this was not the case: 0.34 birds/km^2 vs 0.14 birds in pairs/km^2, a difference of 0.20 birds/km^2. Clearly, the large number of apparently non-breeding dabblers needs to be confirmed or explained. Potential reasons include (1) unusual low breeding propensity, (2) aggregation of pairs at the approach of the aircraft, which led to confusing them with flocks of non-breeders, and (3) timing of the surveys. Flocks en route to more northerly breeding grounds could be mistaken for non-breeders in surveys conducted too early in the season. In turn, non-breeders could be confused with post-breeding groups of males in late surveys (as stated on p. 31 of the report). These results raise questions about the validity of these surveys for dabbling ducks and stress the importance of their timing in relation to the birds' breeding chronology, which in turn varies among species and between years.

The correction factor applied to the 1992 data (x1.3) was based on a 25% significant decrease in mean total waterfowl densities compared to 1991. This overall decrease, however, was mostly accounted for by the decline of only three species (Canada Goose, Green-winged Teal and Scaups) among the eight species recorded (Table 3). Furthermore, the mean density of all other divers (five species) significantly increased from 1991 to 1992. It may be relevant to examine how this discrepancy between species may affect the overall avoidance criteria and its consequence in closing specific areas. Finally, the correction factor was applied to both areas assuming that the 1991 data represented the true density and that the same correction factor applied to both areas. These two assumptions have not been validated.

Plots with above-average densities of total number of birds were recommended for closure. This did not take into account potential differences and levels of vulnerability among waterfowl species. By selecting areas with high potential, a higher mean value is obtained and thus sets the threshold for exclusion of LLF at a higher level, releasing more areas for LLF. This approach is also intrinsically biased because if LLF reduce the density of waterfowl, more and more areas will be open to flights if the same threshold is applied. Four plots in LLF area no. 1 were closed in 1991 but opened in 1992. If these plots did constitute suitable habitat in 1991, they should have been suitable in 1992 too because habitats do not change dramatically between years. Differences may only be due to other variables such as weather or timing of the survey.

The use of mean densities is also questionable. The frequency distribution of densities of all plots should be graphed to see where the mean is located with respect to the median and various percentiles. The unit used (density) may not be appropriate to reflect key habitats supporting high concentrations of waterfowl. A large plot having low density may be more productive (in absolute terms) than a small one with higher density.

General comment

As stated by the authors (p.1), the long term goal of the Waterfowl Avoidance Monitoring Program is to identify key habitats that support high concentrations of waterfowl and to recommend their closure to LLF. Data collected should be examined more thoroughly to substantiate their pertinence and validity in attaining the intended goal.

Progress Report - Monitoring Waterfowl Breeding Densities
by M. C. Bateman and A. H. Hicks. *1997.*

Stated objectives

This is a progress report of a **study designed to test the effect of LLF** on breeding densities of waterfowl.

Observational methods

Surveys were conducted from 1995 to 1997 by helicopter in 5x5 km square plots based on the Black Duck Joint Venture (BDJV) procedures. Dates of the surveys were not mentioned and a difference of only two weeks in the breeding chronology of waterfowl may explain some of the variation observed in the data. It is crucial to precisely time the surveys during the same phase of the breeding cycle of the various species to get comparable annual data.

Sampling design

The sampling design explained in the Methods section is not clear. By looking at Figure 1, we understand that there were nine 10x10 km experimental units, each associated with one of two treatments or constituting a control unit (no overflights). The two treatments included a before and after treatment (B/A) where overflights did not occur in the first year but did in later years and a treatment with an anticipated high frequency (HF) of overflights. In the Results and Discussion section, we can read that there were four 5x5 km subplots (which were their sampling units) nested in the larger plots. This does not represent 12 independent plots per treatment, as alluded to in the Methods section. Statistical treatment differs depending on whether the data were obtained in a nested design or in a completely randomized scheme. Moreover, we do not know if the four subplots of a given plot were surveyed on the same day. If so, total numbers in the four 5x5 km plots could be used in a simpler ANOVA model.

Statistical treatment and data analysis

The authors mentioned that densities of birds were different on plots with the same treatment (p. 4), but there are no statistical tests to support this claim. Based on this assumption, they narrowed their analyses to simple paired t-tests between years for each treatment and control plot. They are not really testing the effects of the LLF, but only the annual variations without consideration for other potential sources of variation such as weather, migration chronology, recruitment, or winter mortality. Usually, a design with controls and treatment plots involves testing the difference between them knowing that other sources of variation affect all experimental units alike. A straightforward ANOVA with specific contrasts would have been interesting, at least to show the variation in the data.

It is unclear how their statistical treatments have considered the nested design in calculating means and standard errors in the paired t-tests (experimental or sampling unit means?).

Presentation and interpretation of the results

Standard errors should have been presented in Table 1 to give an idea of the variability in the data.

If these results are to be compared with similar studies, it is important to clearly define the treatments. Only flights that occurred in 1996 are presented (Table 2). For ease of comparison, plots should have been pooled by treatment category in Table 2. Also, LLF activity was given as 22 April to 31 October: since this study focused on the breeding period, only flights during the April–June period should have been tallied. One could argue that the cumulative effect of the flights (April to October) should be considered, but it is unlikely that the birds potentially affected by the LLF during the molt and fall migration are the same that were present during the breeding period. Lastly, no information is provided about the homogeneity of treatment within the larger (experimental) units and we have to assume that the flights influence equally all portions of the 10x10 km plots: this needs to be confirmed.

One must be careful in interpreting results from paired tests with a factor of more than two levels (three years in this study) because assertions that emerge from a paired comparison may be generalized over the entire set of data. As an example, the authors mentioned that total Black Duck numbers were higher in control plots and on high frequency plots. This is true for the high frequency plots because the 1996 value was significantly different from the 1995 and 1997 values, but it was not so for the control plots. A significant difference was observed only in 1997. The same caution must be applied when reporting differences between years in different treatment plots since no tests were done on treatment effect.

It is not known if all 12 plots were used in the analyses or only those with ≥1 observation. Black Ducks and Canada Geese were the most abundant species, followed by Red-breasted Mergansers and Surf Scoters although the survey period was probably too early for the latter.

General comment

It is an interesting study but data should be re-examined using adequate statistical design and testing, which was the objective of Turner and Hicks (2000).

Breeding Population Trends of Waterfowl in the Military Low Level Training Area
by B. Turner and A. Hicks. *2000.*

Stated objectives

This study tested the effects of LLF on breeding waterfowl by examining population changes over the years. It is essentially the same study design and data presented by Bateman and Hicks (1997) with the addition of three years of survey.

Observational methods

See our review of Bateman and Hicks (1997).

Sampling design

There were two treatments (overflight frequency) and a control. Each level of treatment was assigned to 3 (replicates) 10x10 km plots (experimental units) in the Labrador region, which makes a total of 9 large plots. It does not appear that plots were randomly assigned to a treatment, but this may be assumed as such provided that it is not unreasonable. Data were collected during 5 consecutive years on the same plots.

Statistical treatment and data analysis

By looking at the model presented in the data analysis, it seemed that a split-plot design in time was used to analyze the data, which is the appropriate analysis. Intra-plot variation measures the effect of the year and inter-plot variation estimates the treatment effect. The term *location* in the model represents the treatment effect. This could have been more clearly expressed.

The existence of a linear trend in the numbers of birds over the years was tested using a contrast in the ANOVA. This was appropriate since it has the advantage of using the intra-plot variance instead of the inter-plot variance as in a regression analysis.

The ANOVA was used only to test the difference between high frequency overflight and control plots. Before and after plots could have been included in the ANOVA with proper contrast analyses.

The purpose of using transformed data instead of raw data is to make the means and variances independent and to obtain homogeneous variances. There is some redundancy in presenting results in the tables with and without transformed data. A short discussion on the improvements resulting from the transformation could have been presented and only the results based on the transformed (or raw data if no improvements) should have been presented.

The authors have used an approach based on the power of the tests for evaluating the number of years required to detect a significant trend. Several simulations were done using different slopes (absolute changes in the number of birds). The formula presented on page 8 only described the first step of the simulation. It would have been more helpful had the authors used a notation similar to the one used in the previous model. It is not clear how this formula was applied and which other parameters were used (critical F, degrees of freedom, non-centrality parameter) in calculating the power.

Presentation and interpretation of the results

To ensure that we understood their approach, we repeated their analyses using data provided in Appendix 1. We reproduced the results of their Table 1 with only slight differences for Black Duck despite the fact that a site was removed (DND-9 and not DND-1 as specified in the text). There is probably an error in the raw data or some values considered as outliers were removed from the analysis without this being mentioned.

We were unable to replicate the results of Table 2 even when using a two-way ANOVA as mentioned by the authors (p.22), which is not the same model as initially stated. Using the correct split-plot in time model, we did not obtain the same statistical results either, although we reached the same conclusions. We could have tried to determine the type of analysis used had the authors presented the degrees of freedom associated with the F values.

Viewing trends in terms of a regression analysis is useless since no linear trends were significant in the ANOVAs (Table 1). Calculating R^2 for each plot was somewhat overestimating their usefulness since there were not enough data points to achieve any significant relation. Even the authors clearly stated this fact (p. 7). Fig. 6 to 20 are only useful to show trends within or across treatments (as indicated by the authors in their results section) and to support the ANOVA results. In that respect, only the figures for Black Ducks and Canada Geese are interesting. Density estimates for groups of ducks (dabblers and divers) are too dependent on the timing of migration of the various species (in relation with the survey dates) to draw any conclusion.

We were also unable to reproduce the power analysis simulation presented in Table 3 when using proper degrees of freedom and mean square error derived from the ANOVA. This is not surprising since at first, the overall ANOVA results were not replicable. Still, it is rather surprising to reach a power of 0.95 so quickly in view of the existing variability.

General comment

This report is a major improvement over Bateman and Hicks (1997). Statistical treatment seemed appropriate although the data analysis section does not describe the methodology adequately. Data should be reinvestigated and results presented according to stated ANOVA models.

Waterfowl Behavior in Response to Jet Overflights at Snegamook Lake, Labrador
by M. C. Bateman, A. H. Hicks and S. M. Bowes.*1999.*

Stated objectives

This **study was designed to test the effects of LLF** on the behavior of Black Ducks and Canada Geese during their molt and fall staging between 1995 and 1997.

Observational methods

Behavioral observations were noted using scan sampling at 15-minute intervals from two blinds located at Snegamook Lake. It is not clear how the data from the two blinds were treated (totals or means?). The activity of the birds was recorded using standard categories including feeding, swimming, loafing, preening, and flying. One activity missing in their study is the time spent by birds in alert posture, which should have been included since they were looking at the effect of disturbance. Swimming and flying may be considered in some cases as an escape reaction. The authors also reported an alternate site of observation but they did not specify if data collected there was comparable to that from the other sites.

Sampling design

There were three classification criteria (year, period of year, and time of day) and one treatment (the low level flight of a jet passing over the area). Data collected are percentages of birds in a particular activity. This suggests a three-way ANOVA comparing before and after treatment. Data were collected on a temporal scale, which means that there is necessarily a dependency between observations. This dependency must be addressed properly in any ANOVA and can be sometimes difficult to assess.

Statistical treatment and data analysis

Although an overall ANOVA could have given some insight on the effects of year, period of year, and daytime period interactions with the treatment, the authors have chosen to analyze their data separately for each combination of factors (year x period of year x daytime period). This is a simple approach and probably the best one since for certain combinations there were no data available (missing cells in ANOVA). However, interpreting the results of such multiple tests is difficult.

The data are reliable although the number of observation periods with flights is limited (1996–1997). Because of this, statistical testing using the normal model may be questionable. Total numbers of birds doing a specific activity follow a binomial distribution and will approximate normal distribution when numbers are more than 30 and the percentages are between 20 and 80%. A binomial (logit) model may be more appropriate.

It seems that percentages and their standard errors were calculated using the mean of the percentages calculated for each scan. It is usually preferable to use the total number of birds observed in evaluating ratios and standard deviations (Cochran 1977:145) because it gives a more realistic estimate instead of giving the same weight to each scan that has different numbers of birds.

The ANOVA model and the method used to compare means between years are not mentioned.

The small sample sizes for observation bouts following flights require evaluation of the power of the tests but an alternative approach is suggested. The effects of the flights could have been evaluated with paired t-tests comparing the % of time devoted to various activities during the scan before and after the flights pooling all years, periods of the day and season. This would have reduced the variance because there is a chance that the same flock of birds was present both before and after the flight. Using all bouts with no flights increases the variability and greatly biases the estimate; the objective is not to precisely describe the time budget of these birds in different years, seasons or periods of day but to evaluate the effect of the flights. This would have resulted in a more robust evaluation of the effect of the flights because all flight occurrences would have been used instead of using the year, period of the season, and period of the day combinations with sufficient data. Moreover, by comparing data from the bout before vs the bout after 15 min., then with the bout after 30 min., 45 min., and 60 min., an indication of the persistence or the fading of the effect could have been obtained. Both the number of birds and the percentage of time spent in the different activities could be analyzed this way.

Presentation and interpretation of the results

In general, this report is too long with redundancy in data presentation. Many tables and figures are not essential to document the effects of LLF. Table 1 should be deleted since information in the text is sufficient. Percentages showing overall activity of birds based on all observations (Tables 3, 5, & 6, Fig. 3 to 6) are misleading because the data include, at least technically, a possible treatment effect.

Table 4 was presented to show the importance of Snegamook Lake for waterfowl and to justify its choice for carrying out the study. A graph showing maximum or mean number of birds of each species against day of year would have been more informative than the total numbers of birds summed across all scans. Table 2 should have summarized the treatments (overflight frequency) by presenting data according to the classification used in the statistical analyses. Moreover, the raw data about the overflight frequency (Table 13) should have been presented in an appendix.

As mentioned before, interpreting ANOVA results for each combination of factors must be done with caution (Table 17–22). For instance, significant effects of LLF were found in 4 (25%) of the 16 paired comparisons for Black Ducks. Can we therefore conclude that LLF have no effect?

Results of the effects of year, period of year, and daytime on the controls were presented as such (Tables 7 –12). Examining these in view of the statistical analyses would have been more useful because these factors were considered with the sole purpose of controlling the sources of variation.

General comment

Proper re-analysis of these data is required before any conclusions can be drawn.

Waterfowl Populations in the Low Level Training Area of Labrador–
A Data Compilation and Analysis
by M. C. Bateman and A. H. Hicks. *1999.*

Stated objectives

This report is a compilation of data from all surveys conducted in and around the LLF zone from 1980 to 1998. Stated objectives were to establish whether or not surveys conducted in the past were adequate to provide baseline information in view of implementing measures that would mitigate the potential effects of LLF. **These surveys were not designed to test specific hypotheses about the potential impacts of LLF**.

Methods used in this compilation

The data were compiled by year, month (grouped according to specific periods in the waterfowl life history: spring staging, breeding, molting, and fall staging), and ecoregion (Lopoukhine *et al.* 1977). However, there is no description of how data from different sources were treated to calculate means and standard deviations by ecoregion.

When carrying out such a compilation, one is always confronted with diverse observational methods and sampling designs. The reader must be able to evaluate how comparable the data sets are. A table presenting the main characteristics of the data would have been essential. It could have included reference, survey program, methods (transect or plot, aerial or ground), years and purpose, sampling effort, sampling design, etc.

Presentation and interpretation of the results

The authors first reviewed breeding pair surveys conducted by helicopter in 10x10 km plots. Except for the Smallwood Reservoir (n=15), data for the other ecoregions were characterized by low sample sizes (mean=2.4 plots/ecoregion, range=1–6, Appendix B). It is not known how SD were calculated in Tables 1 and 2 for ecoregions with a single plot (Lake Melville, Porcupine Strand). It is also unknown if the same plots were surveyed in more than one year and how this was taken into account. We suspect that some plots were surveyed in different years and pooled, which is a clear evidence of lack of independence. No years were specified in Tables 1–6. Comparisons among ecoregions would be biased if they were surveyed in different years. Extrapolation of density value from plots or transects should be viewed with caution because some surveys were done in prime habitats with known or suspected concentrations of waterfowl (p. 3).

The whole purpose of conducting surveys is to estimate the number of birds that is statistically representative of an area. This can only be achieved if the surveys are based on an adequate sampling design. One must keep in mind that raw census figures shown on a map for a given location are not very informative since they do not reflect variability (means with standard deviations) in the data. Furthermore, raw data only describe a situation existing for the set of locations surveyed and not for the whole geographic area unless the latter has been totally covered as in satellite imagery, for example. Therefore, maps showing locations of important areas for waterfowl within the LLF zone should be used with caution since the absence of birds at a specific site in a given area does not necessarily mean that it is not important, but could simply indicate that no surveys have been conducted there.

The same argument can be made for maps showing average density of Canada Geese per ecoregion (Fig. 6–9). Aerial transects were flown to get an overall density estimate of the northern landscape and the data were subsequently stratified by ecoregion. The calculated means may not be representative and one must be cautious in interpreting the data. For example, an average density of 5 to 10 indicated pairs of geese was estimated in 1998 for the Hopedale ecoregion located along the Labrador coast (Fig. 7). In Fig. 5, however, we can see that only a very small portion of transects was flown in the southern part of this ecoregion.

General comment

It is difficult to see how the mapping of the area in ecoregions could be useful for waterfowl since these ecoregions are not based on waterfowl habitats. Presenting the numbers of birds per month and year does not help one gain a general perspective on the usefulness of the information in assessing LLF effects.

General Review

Continental Aerial Surveys

Bateman and Hicks (1999:1) clearly stated that one conclusion of the 1994 environmental impact assessment was that “the knowledge of waterfowl in Labrador was less than adequate for waterfowl management as well as for mitigation of any effects of jet overflights.” Based on our review of the documents listed in Table 1, we also conclude that data remain inadequate. The main reason is that most of it comes from aerial surveys conducted to estimate overall numbers of waterfowl in Labrador and their relative contribution to the Atlantic Flyway population. Such aerial surveys are well designed for that purpose, but the sources of variation are numerous and it is easy to argue that the fluctuations in the estimated numbers of birds are caused by LLF. So, general aerial surveys over the territory will always be useless in determining the effects of LLF. There is an urgent need to develop well-designed studies that will test the effects of LLF on waterfowl throughout their annual cycle. The recent studies by Bateman *et al.* (1999) and Turner and Hicks (2000) are improvements along this line but there is still some way to go.

The Concept of Ecoregion

Several reports mentioned Lopoukhine *et al.* (1977) classification of land into 27 regions and used the regions to generate estimates of density. These estimates are always based on *a posteriori* analyses of the data. They can only give an idea of the densities and cannot be used as statistically representative values of each land region. Furthermore, we do not understand the need for such a classification. No studies have shown a real link between waterfowl habitat and the land regions of Lopoukhine *et al.* (1977). We did not review this report but Goudie and Whitman’s (1987) description of eight of these ecoregions showed that each land region includes a variety of habitats suitable to various species of waterfowl. Estimated densities are mostly related to the number and extent of suitable habitats and not the overall ecological features of a land region. A highly prized and unique habitat may be lost or degraded through disturbance only because it is classified in a land region with an overall lower bird density. A real northern waterfowl habitat classification is needed and should be on a per species basis (or at least on the basis of the major species groups) since the ecological requirements are different for each community component. **The ecoregion concept does not reflect the capacity of the habitat to support populations of wildfowl.** The use of GIS tools would be a valuable asset in the future to locate key waterfowl habitats and to assess the potential impact of LLF.

Notes About Measuring Disturbance

Some of the reports we reviewed have quoted Bélanger and Bédard (1989), who studied the effect of disturbances on Snow Geese staging near a small municipal airport (Montmagny) used by propeller planes. It is perilous to extrapolate from a study of disturbances by small aircraft to draw conclusions about disturbances by army jets. The duration of the disturbance associated with LLF lasts a few seconds and noise is probably the most significant factor. However, when we see Snow Geese getting accustomed to propane guns used to scare them off in agricultural fields, we can only wonder about the effect of LLF noise. Before extrapolating results from other studies, we should bear in mind that helicopters and small prop-planes cannot be compared to jets, at least in the case of geese.

There is a need for a long term monitoring program to see whether waterfowl avoid the more severely disturbed areas. Control plots are needed and they should have the same biophysical characteristics as the LLF areas. The survey methods developed within the framework of the BDJV for breeding pair surveys (see Collins 1999) should be used as a basis to establish methodology: helicopter, plot size, survey dates, statistical analyses, etc. Adjustments would be required for late nesters. On the other hand, the BDJV methodology for breeding birds may not be suitable for molting and migrating birds, which disperse in clumped distributions (flocks). In addition, it would be necessary to take into account the turnover rates of individuals during the migration periods in such flocks. It is probably not the same birds that are constantly disturbed at a given site. Concerning reproduction, the only approach is to see whether recruitment in an area influenced by LLF is significantly different from that in control areas. Several years of banding using recent analytical CMR (capture-marking-recapture) procedures would be an interesting approach. The concept of source and sink could also be used to evaluate the impact of the LLF. However, low densities and limited access to most of the area would make such a study difficult and costly. Nevertheless, at least for Black Ducks, it could be a part of the BDJV. Most of the data used to estimate survival and hunting mortality of Black Ducks come from banding operations conducted in the southern range of the species. We do not know if inferences made for these southern populations apply to those of Northern Quebec and Labrador. Coupling local needs (effects of LLF) and continental ones (population dynamics) could justify such a study.

Density as Avoidance Criteria

The concept of a critical density is not easily applicable because of the great variability of this parameter. Factors influencing density of waterfowl include the biophysical characteristics of a site, but also extrinsic parameters such as weather conditions, migration and breeding chronology, population dynamics (e.g., mortality due to hunting occurring further south), and characteristics of the surveys themselves (observers, type of aircraft [helicopter or fixed-wing], flight altitude and speed, correction factors [which must vary according to species], size and distribution of plots, etc). In addition, density is calculated from the number of birds observed divided by the inventoried area. It is never clear how this area is calculated and what proportion of the plot consists of suitable habitat (wetlands), excluding upland and deep water areas. Most of the data appear to represent overall densities as mentioned before. This is adequate to establish population sizes (e.g., BDJV), but not to establish critical thresholds nor to characterize sites or zones.

To cancel flights over plots where waterfowl density exceeds a certain threshold will not be feasible, whether plots are blocks of 10x10 km or vary in size and shape (see RRCS 1993). The entire area will look more like a Gruyere cheese with many small zones where flights would be prohibited in a matrix of less suitable habitats where flights could take place. A landscape approach would be more useful. Landscapes could be characterized according to the habitats suitable for the various groups of wildfowl at various times in the annual cycle. Density could then be evaluated for these various habitats and guidelines established to close entire zones where there is an abundance of suitable habitats. These densities should be established for the various phases of the annual cycle considering the overlapping chronology of the various species.

Perspectives and Future Issues

Testing LLF Effects on Waterfowl

The primary goal of the program is to determine whether or not the LLTF (treatment) impact on waterfowl numbers and distribution, and at what level of flight activity such impacts become prejudicial to the bird community; impacts should be detected using parameters such as presence/absence, trends in abundance, breeding output, length of the staging period, shifts in migration routes, time devoted to feeding, etc. The ultimate goal is to ban LLTF in areas where waterfowl is concentrated or to maintain flight frequency therein below the threshold at which impacts were demonstrated on one or several of the parameters.

The problem sounds simple but is eminently complex. First, we will never know to what extent current waterfowl distribution has already been influenced by several years of training flights covering most of the study area. Some or several of the waterfowl species may already have adapted their distribution and abundance in response to past flights and trying to measure an impact now may be futile. Second, the waterfowl community itself is quite diverse and the various species involved are unlikely to present a homogeneous response to the treatment; for instance, Snow Geese will be panic-stricken while several dabblers will sit passively during the passage of a low flying conventional aircraft and similar divergences may exist among the Labrador waterfowl species assemblage in response to jet overflights. Ensuring consistent treatment application is also in itself an enormous challenge: treatment level varies with aircraft type, altitude, proximity, and approach angle. Response to the treatment by the various species may also vary according to the annual cycle of the birds, making it difficult to establish paired control and experimental plots that are comparable in every respect[1]. Furthermore, detecting and measuring a behavioral response to an overflight is NOT the same as detecting an impact. Alertness to environmental stimuli (the response) is predictable: it is only once stimuli exceed a certain threshold that negative, measurable physiological effects are induced and measuring this threshold is a very difficult objective. So, studies revealing a response will need to be interpreted with caution.

Responses of the birds to the treatment will not necessarily be easily measured as they may consist of numerous, sometimes subtle and most of the time complex and interrelated behavioral changes; it is unlikely that a single-phase, easy-to-monitor response will express itself as several parameters could be affected simultaneously. Breeding performance and distribution during staging may be affected in a linked manner. For instance, changes in distribution, in faithfulness to a staging/molting/ breeding areas, in pair-bond stability (frequency of nest abandonment), and in breeding output (nest losses, brood losses) may all be affected at different thresholds. Such

[1] For instance, control and experimental plots in the study by Bateman and Hicks differed among themselves right from the start, which made it impossible to reach firm conclusions.

parameters are difficult to monitor in the wild with any degree of accuracy in the best of circumstances, let alone in those that prevail in the vast Labrador wilderness. The impacts, should they be present, may also only become apparent after a considerable (two, three, five years?) time lag. Finally, many of the responses observed in bird distribution and abundance could result from events occurring outside of the area and not from the treatment itself. In short, no experimental design will make it possible to untangle local treatment effects from regional or continental ones. To conclude that the problem is untreatable is probably close to the truth and we believe that a convincing and flawless overall experimental design enabling one to reveal effect(s) of LLTF on waterfowl is probably impossible to set up.

In this context, we favor a very different approach to the problem of mitigating impacts of these flights upon the waterfowl community. Using existing knowledge about the requirements of the various species in terms of habitat requirements and timing of the breeding season, it would probably be possible to pinpoint sensitive areas that should always be excluded from overflights without developing an experimental program whose outcome is questionable from the start. A precautionary approach in this context seems appropriate to us and is legitimate in view of the existing body of knowledge about the response of these animals to disturbance.

Defining Key Habitats

As stated earlier, the main objective is to move flight activities to areas where the likelihood of disturbance to waterfowl is minimized. Defining such areas on the basis of bird numbers alone will always yield large variations in numbers depending on year, date of survey, observer competence, weather during surveys, etc. Furthermore, small plots delineated only by bird numbers do not reflect ecosystem integrity based on the ecological requirements of these animals. The approach we propose bypasses this difficulty: Instead of attempting to detect one or several responses among selected species, we rather propose to look at overall habitat characteristics. Existing knowledge about the requirements of the key species is probably sufficient to predict current waterfowl distribution and pinpoint sensitive areas. Ecological mapping of key waterfowl areas must come first and in our view is essential and can be made using satellite imagery and topographic maps. The framework of key habitats for each species would become the basis for managing the closure of areas to overflights. Supplementing this information with aboriginal knowledge and fresh field data when needed would probably be more than sufficient to define a truly precautionary approach. Spatial analysis can then be done, fusing small patches into larger blocks of habitats in which several species are concentrated. Large habitat patches would allow the maintenance of ecological functions essential to each individual species; intertwined with patches essential for the maintenance of similar habitat patches for the protection of ungulates and human requirements would produce a complex grid over which LLTF scheduling for DND purposes would take place.

Confirming Responses of Waterfowl to LLTF

We do not dismiss the need for specific, well-designed experimental studies on key parameters, including productivity/recruitment in breeding areas and time spent foraging during staging/molting by several or all species. Such studies, no matter how academically exciting and valuable they may be, will not in themselves guarantee the maintenance of a population, nor will they help determine the capacity of the habitat to support populations of wildfowl. For this reason, the general precautionary approach seems essential.

The Long Term Uncertainty

When dealing with conservation issues, one is always confronted with the question of long term effects. This can only be answered through long term monitoring and should be specific to certain species at critical periods of their annual cycle. This can be done only for the most abundant species (e.g., Black Ducks and Canada Geese). The Black Duck Joint Venture program already implements such a long term design; but the BDJV objective is to reflect continental population changes rather than local changes. It must be adapted to reflect use of the area subject to overflights. Linking the results of possible impacts of LLTF on waterfowl with concurrent impacts on caribou, osprey, and human occupation and activities adds several layers of complexity to the overall problem.

How to deal with threatened species such as the Harlequin Duck and Barrow's Goldeneyes remains a problem. Designing studies on the effects of overflights is a major challenge as the latter two species occupy the entire range, but at very low density. In the event that LLTF have negative impacts, should we be more concerned about such overabundant and cosmopolitan species as Canada Goose and Black Duck (the consequences of which will be rather minimal on a continental scale) or should we rather be concerned about the fate of the Harlequin Duck and Barrow’s Goldeneye since the suspected impact occurs in the very heart of their only known breeding range?

In Summary

We have concluded that a truly experimental overall approach is not possible for several reasons and we favor a precautionary and less costly attitude that must begin with mapping waterfowl habitat throughout the entire low level training area (LLTA). Surprisingly, this information does not seem to be available, either for the dominant species or for the rare ones of special concern. This mapping should concentrate on breeding, but also on staging areas; it should yield sufficient information to determine a priori avoidance flight zones, notwithstanding whether or not effects are occurring. This process should include the following steps:

1) Determine where the various species are found (fast waters, shallow acidic lakes, water bodies with floating emergents, marshy areas, etc.) based on available (published) knowledge and standard geographical tools (satellite/conventional photos).
2) Run field tests to ascertain whether predicted quality areas are in fact used by waterfowl and obtain concurrent density estimates. During this process, determine a convenient and amenable plot size to obtaining density measures in a rapid and reliable manner.
3) Rank the various species in terms of importance/vulnerability and decide where emphasis will be placed for specific studies.
4) Be prepared to take regular measurements over several years (5 to 10 yr.) before expecting density changes that can be interpreted. Specific studies showing the evolution of selected parameters (nest abandonment, survival of ducklings/ productivity, behavioral changes in alertness, flight behavior, etc. must be carefully planned and must use an experimental design. Such studies are costly and it is notoriously difficult to obtain a reasonable sample size.
5) Lastly, the various reports we examined lacked in uniformity of approach and methodology and were not consistent in the goals pursued. To be successful, the program we recommend must be planned and closely supervised throughout by a single expert committee so as to ensure cohesiveness and unity of purpose. Individual studies could be contracted out to consulting experts and/or students enrolled in various graduate programs, but contrary to what has occurred so far, these studies must be carefully integrated through the supervision of a single coordinator. Design of the individual studies should also be assessed by independent reviewers before being undertaken while the entire program should be periodically reviewed by external referees. Once more, emphasis should be placed on studies that will eventually reach the primary peer-reviewed scientific literature. This will give much greater credibility to the studies and monitoring programs. Most of the government or consultants' reports that we have reviewed would not reach this level. Good examples of applied research that translated into scientific papers and that were used as part of EIS are those involving the Spotted Owl on the West Coast (ex. Call *et al.* 1992, Lamberson *et al.* 1992, Murphy and Noon 1992).

Acknowledgements

We thank Mr. Gaëtan Daigle of the Service de consultation statistique, Département de mathématiques et de statistiques, Université Laval for his well-thought-out guidance and advice in evaluating the statistical methods used in some of the documents reviewed.

References

Bélanger, L. and J. Bédard. 1989. Responses of staging Greater Snow Geese to human disturbance. J. Wildl. Manage. 53(3): 713–719.

Call, D. R., R. J. Gutierrez, and J. Verner. 1992. Foraging Habitat and Home-Range Characteristics of California Spotted Owls in the Sierra Nevada. Condor 94: 880–888.

Cochran, W. G. 1977. Sampling techniques. John Wiley and Sons, New York, ix + 234 p.

Lamberson, R. H., R. Mckelvey, B. R. Noon and C. Voss. 1992. A Dynamic Analysis of Northern Spotted Owl Viability in a Fragmented Forest Landscape. Conservation Biology 6: 505–512.

Link, W. A. and J. R. Sauer. 1994. Estimating equations estimates of trends. Bird Populations, 2: 23–32.

Lopoukhine, N., N. A. Prout and H. E. Hirvonen. 1977. Ecological land classification of Labrador: a reconnaissance. Fish. and Environ. Can. Ecol. Land Classif. Ser. No 4, 85 pp.

Murphy, D. D. and B. R. Noon. 1992. Integrating scientific methods with habitat conservation planning: reserve design for northern spotted owls. Ecological Applications 2: 3–17.