EVALUATING THE QUALITY OF CANADA'S 1986 CENSUS OF AGRICULTURE

M. March, J. Frederick, E. Wilson and D. Castonguay
Statistics Canada

1. INTRODUCTION

The Canadian Census of Agriculture has been carried out for more than a century; every 10 years from 1871 to 1951 and every five years since 1951. Data are collected for a number of variables such as crops, livestock, farm land, labour and capital.

The 1986 Census of Agriculture was collected simultaneously with the Census of Population. Census takers were required to identify all farm operators and the corresponding farmland as they canvassed their areas for the Census of Population. If anyone in the household operated an agricultural holding (defined as having sales of agricultural products of at least $250 in the past year) they were asked to complete a Census of Agriculture questionnaire.

Completed questionnaires were edited by the Census representatives and forwarded to Statistics Canada's Head Office in Ottawa where the information was captured, edited, adjusted for non-response, validated by subject matter specialists, subjected to confidentiality protection procedures and then released in summary form.

The Census of Agriculture provides users with extensive data on farming. Given its size and complexity, however, it must be recognized that errors will occur. Coverage and response errors are almost inevitable. Coverage errors occur when a farm is missed or counted more than once, or a holding is included which falls outside the agricultural universe. Response errors include misinterpretation of a question by a respondent, an error made by a proxy respondent, partial or total non-response, errors in field editing and follow-up and processing errors made during handling, data capture, coding, editing, imputation, etc.

Major users of Census data include: Agriculture Canada (Canada's federal department of agriculture), provincial governments, the private sector and educators. Agriculture Canada requires the data for policy planning, as a base for subsidy programs, to evaluate the impact of past policies and programs, and to identify geographical areas that require assistance. Provincial governments use the data to determine the health of the agricultural sector of their economies. The private sector requires good small area data for planning agricultural marketing programs, production runs and plant locations and for making financial decisions. Educators exploit it for research and analysis. All demand a good assessment of the quality of the estimates.

The purposes of the data quality evaluation are to provide users with a quality assessment of the data that will permit them to interpret and use the information correctly and to provide Statistics Canada staff with evaluation results to be used in improving the design of the Census of Agriculture. Problems need to be identified and quantified; ways of solving them need to be specified for future censuses.

This paper describes the evaluations of the 1986 Census of Agriculture data quality that were undertaken and the measures that were obtained from them.

2. THE EVALUATION APPROACH

The data quality evaluation embodies two components, an evaluation of the Census process and an evaluation of its final product, the estimates.
Evaluation of the Census processes is an important part of assessment of Census quality. In theory, the Census process covers the carrying out of the Census from the design stage and then drop-off of the questionnaires through to the final compilation of summary estimates. Ideally, the entire process should be evaluated from start to finish.

The Census final product is the set of tables, data files and publications that are produced after processing the collected data. Evaluation of the product consists of the analytical examination of the estimates and comparisons with other sources. Both macro aggregates and micro-level data are analyzed.

Both approaches are used because an evaluation of the product alone, while indicating possible errors, does not usually provide insight into their causes. Also, where comparisons indicate inconsistencies, it is not always possible to know whether information from the other source is necessarily more correct than the Census information.

This paper discusses methods used and results obtained through both approaches. Discussion of the product evaluation approach has been mostly limited to the comparisons of the Census with the National Farm Survey, an annual agricultural survey at Statistics Canada.

3. EVALUATION OF THE PROCESSES

The Census of Agriculture was carried out as a series of five tasks that were referred to as data collection, data capture, editing, imputation, data validation and outputs. Where possible, each of the processes that made up a task was evaluated.

A. Data Collection

Evaluation of collection of Census data was limited to computation of the response rate (more than 99% of operators identified completed a form), a study of forms completed by Census representatives to document refusals and an analysis of field quality control data. The study of refusal record forms indicated that both partial and total non-response was most frequently due to a reluctance of farm operators to provide financial data. Analysis of quality control data revealed that about 13% of questionnaires contained at least one edit failure.

B. Data Capture

The data capture process consisted of two operations: document preparation and data entry.

In the document preparation operation, each questionnaire was prepared for data entry; adjustments were made (where necessary) to the answers provided by the respondents to ensure that the data were clear, readable and complete. Consistency checks and corrections for non-response were left to be handled after the completion of data capture.

The data entry operation consisted of converting the data on the questionnaire to a machine readable format; all of the data on each questionnaire were thus key-entered on terminals exactly as reported by the respondents or as clarified by the document preparation staff. The keying program included on-line edits and 100% verification of selected fields.

The on-line edits consisted of basic checks of key-entered values against maximum possible values and previously-entered values that should have been consistent. The program required the operator to revise questionable entries before going on to the next screen.

A 100% verification was performed only on selected fields. Data were entered twice using another operator the second time. The second operator was required to resolve any discrepancies that were found between the entries of the same field.

A quality evaluation was undertaken to measure the levels of errors associated with each of the two operations.

The document preparation operation was subject to quality control using a dependent sample verification plan. Quality control records that had been kept during this operation were used to estimate "error" rates where each "error" had been charged because of the omission or mis-application of a procedure. At first verification, 3.8% of the batches had been rejected. On the basis of the data recorded for each batch in determining if it was to be rejected it was estimated that approximately 6.2% of the questionnaires must have contained at least one error. At second review of the rejected batches, less than 1% of the batches had been rejected. The estimated error rate for these batches was 3.9% of the questionnaires with at least one error.

The estimated average outgoing quality (which was the average percentage of questionnaires that contain at least one error coming out of the operation) was estimated at 5.3%.

The quality of the data entry operation was assessed by re-entering an evaluation sample of the questionnaires and by comparing the re-entered data to the data from the original capture. The procedure follows.

A sample of questionnaires was selected systematically in each province, using a random starting point and a selection interval. The sampled questionnaires were keyed again using the same data capture system including on-line edits and and 100% verification of selected fields. This recapture operation was carried out shortly after the original capture and using the same group of operators. The recaptured sampled data were matched against the originally captured data and all the discrepancies were verified.

The overall error rate was estimated at 0.7%. Fields subject to on-line edits and 100% verification were found to have error rates of 0.6% and 0.2% respectively. For fields subjected neither to on-line edits nor to 100% verification, the error rate was estimated at 3.6%.

About 51% of the errors found were due to keying of a wrong value while 30% were due to overlooked values (a misspelled alphanumeric value was considered to be a 'wrong' value). Another 12% of the errors had been captured values where the field had been blank on the questionnaire.

Overall, it was felt that the data capture system had introduced minimal error into the data.

C. Clerical and Professional Edits

Once Census of Agriculture data had been entered, the data records were passed through a Computer edit system that checked for inconsistencies and identified some of the larger values that had been captured. The editing was carried out in stages. In the first stage, records were run through an editing program which identified fields that contained unexpected values. Using output from this program, clerical staff checked back to the questionnaires to determine if the edit failures had been caused by data capture errors. Where this was the case, records were modified to contain the values actually found on the questionnaires. If the edit failure was not caused by a data capture error, the field was left unchanged. Once clerical editing was completed, the data records were run through the edit system again using a more complex set of "subject-matter" edits. Resulting edit failures were reviewed this time by professional staff who either accepted the values, entered a code signifying that a replacement value was to be imputed or corrected the errors on the basis of subject-matter knowledge or occasionally after a telephone call to the farm operator who had completed the questionnaire.

Since the intention of the clerical and professional edit processes was the improvement of data quality, evaluation of them focused on their impact. Measurement of the human error component was also possible for the clerical edit correction process because it had been subject to a sample verification procedure. The impact of both edit processes was assessed by using information on before- and after- files that was used to compute the impact of each edit stage on every variable that had been processed. The impact was measured in terms of both the numbers of records changed and the total effect of the changes on provincial estimates.

From analysis of Quality control data, it was determined that 7% of the corrected clerical edit batches had been rejected at first review and that failed edits had been handled erroneously approximately 1.4% of the time. Obviously, clerical editing was rarely subject to error. This was not surprising since the procedures had been greatly simplified relative to the corresponding operation in 1981 which had included much more decision-making.

The amount of impact that each of the editing processes had on the data varied from variable to variable. Table 1 illustrates the percent impact of each editing stage on Cattle estimates for Alberta.

Although the impact of the clerical edits was relatively large, in most cases, changes were due to incorrect data entry of too many digits on occasional records which had led to creation of large outliers which were now being detected and removed. A very small number of records were affected (for example, 0.3% in the case of beef heifers).

A change at the subject-matter edit stage was usually due to response problems on the questionnaire. For example, in some cases, farm operators had not provided a breakdown by type of cattle but had instead entered their Total Cattle number in the first space of the Cattle section of the questionnaire. The error was

Table 1
The Impact of Clerical and Professional Editing on Alberta Cattle Data

Variable	Clerical	Subject Matter	Final Value
	%	%	
Bulls	-51.3	-6.9	78,996
Milk cows	-25.0	-7.1	124,204
Beef cows	-26.6	+0.6	1,321,556
Dairy Heifers	-11.7	-6.0	50,381
Beef Heifers	-56.5	-2.4	235,790
Slaughter H.	-35.2	+2.3	284,987
Steers	-46.2	11.5	478,147
Calves	-27.7	+1.5	1,253,082
Total Cattle	-9.8	-1.4	3,827,143
Purebred	-52.5	-9.3	287,314

easily detected by professional edits since the first space was intended for the Bulls total- usually a low figure. Understandably, removal of these errors, would have a significant effect on the relatively small bulls estimate.

Table 2 gives corresponding percentages of records affected for the variables given in Table 1.

Table 2
The Impact of Clerical and Professional Editing on Alberta Cattle Data In Terms of the Numbers of Records Changed

Variable	Clerical	Subject Matter	Final Value
	%	%	
Bulls	0.4	0.6	24,571
Milk cows	0.6	3.5	5,574
Beef cows	0.3	1.0	27,655
Dairy Heifers	0.4	4.9	2,689
Beef Heifers	0.3	1.1	17,869
Slaughter H.	0.6	1.7	6,862
Steers	0.5	0.8	13,189
Calves	0.4	1.8	29,513
Total Cattle	3.5	1.7	33,498
Purebred	5.1	1.0	13,202

D. Imputation

Since the 1981 Census, "nearest-neighbour" donor imputations have been used to adjust for item- and record-level nonresponse. For inconsistencies within records, either deterministic or donor imputations have been used depending on the nature and the extent of the problem. The system that carries out these imputations is large and complex.

To some extent, the appropriateness of the imputations that were made was evaluated by the professional staff who examined imputed records that were major contributors to the Census estimates as part of their analyses of the data before release. They noted a few problems at that time. A planned evaluation of matching variables is to be carried out prior to the development of the 1991 system.

On the other hand, the impact of the imputations has been quantified using before- and after-imputation data.

Table 3 is a sample of some of the impact measures that were obtained for donor imputations.

E. Data Validation

The final processing step before the final database was created was data validation. Analysts compared aggregated data with historical estimates, estimates

Table 3
Impact of Donor Imputations on Ontario Cattle Data

Variable	Records	Aggregate	Final Value
	%	%	
Bulls	1.2	+1.1	253,840
Milk cows	1.0	+1.1	424,326
Beef cows	1.2	+1.1	25,677
Dairy Heifers	0.9	+1.1	571,231
Beef Heifers	1.1	+0.9	475,223
Slaughter H.	1.0	+1.0	343,946
Steers	1.0	+1.1	475,199
Calves	1.1	+1.1	77,008
Total Cattle	0.9	+0.8	2,441,785
Purebred	1.3	-0.2	270,534

obtained from recent surveys, estimates provided by contacts in provincial government departments and with information available through Marketing Boards and associations of producers. They also examined data of the largest contributors to the estimates to ensure the apparent validity of their information. Lists of large producers available from Marketing boards were checked against the Census file to ensure that none of them had been missed in the Census.

During this process, analysts found some additional errors and made corrections to the data as necessary. For completeness, this process was also evaluated from the point of view of its impact. Table 4 contains measures of the impact of data validation on cattle data for Quebec.

Table 4
Impact of Data Validation on Quebec Cattle Data

Variable	Records	Aggregate	Final Value
	%	%	
Bulls	0.2	-2.4	27,017
Milk cows	0.2	+0.0	577,743
Beef cows	1.3	-1.3	163,090
Dairy Heifers	0.5	+0.3	252,878
Beef Heifers	0.3	-1.2	38,359
Slaughter H.	0.7	+1.1	21,747
Steers	4.1	-5.1	68,914
Calves	1.1	+0.9	375,834
Total Cattle	0.3	-0.2	1,525,582
Purebred	0.0	+0.0	241,742

F. Outputs

The production of outputs was generally a fully automated and straightforward process except in the application of confidentiality protection procedures to the final tables. These procedures consisted of the suppression of any cells in tables to be published where it had been determined that there was a risk of disclosure of the data for a specific farm holding.

For certain tables, especially those containing small area estimates or estimates for rare commodities, the completeness and therefore the quality of the tables was significantly affected by the confidentiality procedures.

Evaluations of the confidentiality process consisted of a study of the impact of the confidentiality procedures in terms of the frequency with which cells were being suppressed as well as a study of the effectiveness of the procedures used in preventing disclosures.

4. EVALUATION OF THE PRODUCT

The National Farm Survey (N.F.S.) was used extensively in the evaluation of the Census. It is an annual multi-purpose probability survey which was conducted for the first time in 1983. It replaced two other multi-purpose surveys, the Farm Enumerative Survey (F.E.S.) and the Agriculture Enumerative Survey (A.E.S.), which had originated in 1971 as a post-censal evaluation survey. The primary objective of the N.F.S. is to provide estimates of level and change for a wide range of land use, livestock and farm operating expense items.

The N.F.S. was designed in 1983 using the 1981 Census of Agriculture data base and a methodology involving a multiple frame sampling technique. Due to its integrated method of data collection, the N.F.S. can be divided into two components. The first component, the CORE survey, consists of an area frame sample as well as a sample of farms from a list of large operations. The area sample data is collected by personal interview while a mail-out and call-back or pick-up method is used for the list sample. The primary objective of the CORE survey is to provide reliable estimates at the provincial level for all survey items. The second component, the TEL. (telephone) survey, collects data for an additional sample of farms from an extended version of the CORE list frame. A shorter version of the survey questionnaire, containing land use and livestock items but only a few financial items, is administered by telephone. The TEL. list sample, in conjunction with common portions of both the CORE survey area and list samples, is used to produce estimates for sub-provincial areas (SPA'S) as well as at the provincial level for those items on the shorter questionnaire.

Since the 1986 National Farm Survey was conducted in July, only one month after the Census of Agriculture, and covered the same major land use, livestock and farm operating expense items, it was used as a source of alternate, independently-collected data with which to compare the census. However it was recognized that there might be limitations to its usefulness, due to, among other things, slight differences in concept definitions and questionnaire wording, higher non-response rates and different processing systems. The different reference dates, June 3 for the census and July 1 for the N.F.S., were assumed to have minimal effects on the comparison, although differences in responses to inventory questions were expected.

The target population defined by the N.F.S. was identical to that of the census. The population actually surveyed, however, differed from the census in that it excluded: all farms in Newfoundland; all farms in the Maritime provinces, Quebec and Ontario which had received less than $1000 from the sale of agricultural products during the twelve months preceding the survey date and all farms in the Prairie provinces and British Columbia which had received less than $2000 from the sale of agricultural products during the twelve months preceding the survey date; farms located on Indian Reserves; institutional farms; community pastures in the Maritime provinces, Quebec and Ontario; and farms in marginal areas with little agricultural activity. These exclusions, based on 1981 Census of Agriculture data, allowed more efficient use of the resources available for data collection. The estimates produced from the survey data were adjusted to account for these exclusions so that the final estimates would relate to the target population.

The N.F.S. estimates were subject to some of the same types of coverage and response errors as those from the census, although it was expected that these types of errors would have a lesser impact on the survey due to its concentration on a smaller number of holdings. In addition, since the N.F.S. estimates were derived from a random sample. Thus sampling error had to be taken into account when comparing the census and survey estimates. Impact of this survey sampling error on survey estimates for small geographic areas or relatively uncommon items was significant. It was difficult to isolate or measure the effects of non-sampling error on the two sets of estimates.

Comparisons were carried out at macro and micro levels.

For the macro-comparisons, the 1986 Census and N.F.S. estimates for two different universes were compared. The N.F.S. estimates for the target universe (all farms) for all provinces other than Newfoundland were compared to corresponding Census estimates. Then, the unadjusted N.F.S. estimates which corresponded to the surveyed universe were compared to Census estimates for approximately the same universe. (Since boundaries of the surveyed universe had been defined in terms of 1981 Census geographic boundaries while Census data had been coded according to 1986 boundaries, it was impossible to make an exact comparison.) The second set of comparisons was carried out because of a realization that differences between estimates for the complete farm universe might be due to the adjustment rather than to errors made in the collection of the data.

For each macro-level estimate that was compared, a difference between the Census and the survey relative to the survey value was computed as follows:

$$\frac{(C - S)}{S} \times 100\%$$

where C and S were Census and survey values respectively.

Table 5 illustrates the results of some comparisons of data for field crops in Ontario and for the estimated number of farms. N.F.S. coefficients of variation (C.V.s) have been included to illustrate the extent to which sampling error is a problem in the comparisons.

Table 5
Comparison of Census and N.F.S. Ontario Field Crop Estimates (in acres) Target Universe

Variable	Census	N.F.S.	C.V.	Difference
			%	%
Spring Wheat	75,704	70,754	14	7
Corn for Grain	1,829,220	1,983,291	6	-8
Tobacco	64,687	51,663	21	25
Soybeans	939,738	999,987	10	-6
Barley	602,515	622,705	6	-3

Clearly, it is difficult to determine with this comparison alone whether there are problems in Ontario due to coverage or systematic response error because of the relatively large sampling errors. Inconsistencies between the N.F.S. and the Census in the collection and processing of responses also may have contributed to the difference. Certainly, the

tendency to negative values implies some under-coverage but it is difficult to quantify it. Consistent values in the two surveys are reassuring. Inconsistent values suggest the need for further investigation. For these reasons, micro-comparisons were also necessary.

The micro-match between the operations surveyed in the National Farm Survey and enumerated in the Census of Agriculture was performed in two stages. In the first stage, the operations were linked by a computer match, making use of the name and address of the operator. All the operations that could not be matched by computer proceeded to the next stage, which consisted of a manual match. This extra step allowed finding of operations under different names or spellings. Three tools were available for use in the manual search: a list of names and addresses created using the Census file, the Census Representatives' records of visits to dwellings (supplied by the Census of Population) and a search capacity on the Census of Agriculture processing database that was part of the processing system and enabled the retrieval of records with specific characteristics or identification information.

After these two stages, all the operations that were linked were considered to be only 'tentative' matches; in fact, it was possible in some cases to find many 'likely candidates' in the Census to match one N.F.S. operation. All of the matches thus were passed through a validation program designed to eliminate mismatches.

This validation program consisted of a comparison between selected variables reported in the two survey questionnaires. All the variables for which the Census value was not within a pre-defined range of the N.F.S. value were flagged; the matches were then rejected or accepted as valid according to their number of flagged variables.

Table 6 gives matched rates for the micro-match by province for N.F.S. records linked to the Census.

Further analysis has led to the conclusion that the matched rates are not a measure of Census coverage because of the frequency of inconsistent reporting of the operator name and address (partnerships appear under different names, two farms in one survey might be a partnership in the other, the same person may use a different address on a different occasion, etc.).

Table 6

Province	Matched Rate
	%
Prince Edward Island	96.6
Nova Scotia	87.8
New Brunswick	90.3
Quebec	90.5
Ontario	88.8
Manitoba	92.5
Saskatchewan	91.0
Alberta	90.4
British Columbia	82.3
Canada Level	89.9

However, study of linked data can prove useful in the study of both response and processing errors but because of the time and expense involved and the complexity of the analysis it has proceeded only as resources for it have been available.

Professional staff have also used estimates from sources other than the National Farm Survey in evaluating the Census estimates. At the Macro-level, it is reassuring if the estimates are consistent but generally the quality of the information from the alternate source is not well enough known in order to use differences in quantifying error. Clearly, for financial data, significantly higher or lower estimates from Tax data could indicate over- or under-reporting or over- or under-coverage in the Census. Analysts have looked at large differences quite carefully using subject-matter knowledge and for some of them they have been able to attribute a cause. In other cases, there is no explanation.

5. CONCLUSIONS AND PLANS FOR 1991

The two approaches to the evaluation, that is evaluation of the processes and evaluation of the product have led to the gathering of information for data users and planners of both the processes and the evaluation of the 1991 Census.

The process evaluations were useful in determining problem areas in the design of the Census and in the systems and procedures used to derive estimates from it. Results of the process evaluations helped both in the assessment of quality of the data and in identifying problems that will have to be resolved in the design of the questionnaire for the 1991 Census and in the planning of its processing operations for that Census. Clearly the clerical edits are important because of the huge impact that data entry errors can have on the estimates. The impact reports have also identified several questions (or it may be the questionnaire format in some cases) where a re-design should be considered.

The product evaluation identified several situations where there was a possibility of significant under-coverage or problems in the design of the questionnaire.

In planning for the evaluation of the 1991 Census, several improvements are being proposed.

Attempts will be made to obtain data to be used in evaluation of the data collection process. Although it was possible to evaluate the Head Office processes, very little information is available about errors made in the field and about the impact of field procedures on data quality. For example, there is very little information about initial response rates and the amount of editing and correction contributed by the Census enumerator. One objective in planning the 1991 evaluation will be to extend it to include field processes.

It is also possible that use of an evaluation sample might be extended to evaluation of more than the data-entry process. Research is currently under-way at Statistics Canada to generally evaluate the feasibility of using such a sample to evaluate all or most stages of a survey.

A third concern is with the usefulness of the National Farm Survey in the evaluation of Census coverage. Users have indicated that they would like more than the "indication" of coverage available from the macro-comparisons between the Census and the survey. Discussions are already under way to determine ways in which the linkage between the two surveys might be improved to enable reasonably reliable estimates of under-coverage. Possibilities being considered are collection of additional matching information in the survey or contacting the respondent to help in reconciling non-matches.