

## MONITORING SURVEY PROCESSES OF THE CANADIAN MONTHLY WHOLESALE AND RETAIL TRADE SURVEY

Mark Majkowski, H  l  ne B  rard, Catherine Dufour and Susie Fortier, Statistics Canada  
 Mark Majkowski, Business Survey Methods Division, Statistics Canada, Ottawa, Canada K1A 0T6

**Keywords:** monitoring, diagnostic tools, level estimates, trend.

### Abstract

The Canadian Monthly Wholesale and Retail Trade Survey (MWRTS) produces monthly estimates for sales and inventories at various province and industry levels. The MWRTS has recently been redesigned, in part, to provide estimates for the new North American Industry Classification System (NAICS) and to take full advantage of administrative data from the Goods and Services Tax program. The redesign also addressed the need to maintain the quality of the estimates, to reduce cost and respondent burden, to update computer systems, and to harmonise concepts and methods with the annual survey. As part of the redesign, different tools were developed to ensure proper monitoring of survey steps. The first type of 'diagnostic' tool aims to assess the functionality of the modules in each survey step while the second type involves monthly descriptive statistics, such as number of live units in the sample, imputation methods used and recurrent top contributors. These statistics are studied longitudinally to detect changes that affect the estimates. A third set of diagnostic tools is used for analysis of level and trend estimates. Finally, some of the major tools used on a monthly basis for this survey will be described in this paper.

### 1. INTRODUCTION

The Monthly Wholesale and Retail Trade Survey (MWRTS) produces monthly estimates of total sales and inventories for industry trade group (TG) and geographic region (GEO) domains in the Canadian wholesale and retail trade sectors. The resulting estimates for this survey are an important economic indicator since their contribution to the Gross Domestic Product (in terms of value-added) is approximately 12%.

The MWRTS has recently undergone a full-scale redesign to meet a number of objectives. A post-mortem of how well the redesigned survey has initially performed with respect to the objectives is summarised in Tr  panier (2004). The majority of the survey procedures have been modified / updated and this has resulted in a number of monitoring processes

being created in order to ensure the highest quality possible. Section 2 of this paper describes the objectives of the redesign, what was redesigned and the present state of the project. In section 3, some of the major monitoring or 'diagnostic' tools are discussed and explanations are given of why they are useful. Some numerical and graphical examples of these tools are also provided. Section 4 gives a description of how these tools are used together in production and the paper ends with some general conclusions in section 5.

### 2. MWRTS REDESIGN

#### 2.1. Objectives and redesigned items

The MWRTS has recently been redesigned to meet a number of objectives that include:

- a) providing estimates that correspond to the new North American Industry Classification System (NAICS),
- b) taking full advantage of administrative data such as the Goods and Services Tax (GST) data,
- c) addressing the need to maintain the quality of estimates with the reduction in sample size,
- d) reducing respondent burden where possible,
- e) facilitating the collection, processing and releasing of data by moving away from the mainframe environment to one that is pc-based, and
- f) harmonising concepts with the annual wholesale and retail surveys at Statistics Canada that were recently redesigned as part of the Project to Improve Provincial Economic Statistics (PIPES).

The previous full-scale redesign of the MWRTS was completed in 1988. From the period of this last redesign in 1988 to the period of the current redesign that began in year 2000, there have been significant developments in computer technology, industrial classification and survey methodology. This latest redesign is a timely one since these developments will result in a new survey that is an improvement over the former survey. Furthermore, significant problems will be avoided in the future by implementing procedures whose purpose is to maintain the quality of the new survey. Examples of the survey-related items that were modified in the latest redesign included the following:

- a) Made better use of administrative data (that is more up-to-date and expanded for units that are part of the survey frame) in order to develop an efficient sampling plan, to improve the imputation system, and to produce reliable estimates. See Bérard (2001) for more information.
- b) Surveyed non-employer units for the first time. For more information, see Bérard (2001).
- c) Utilised Statistics Canada's Generalized Sampling System (GSAM) to update the sample each month from a stratified design for each TG by GEO grouping (cell) through the addition of a sample of birth units.
- d) The stratification changed to include a take-none stratum within each cell. The take-none stratum consists of the smallest units that are not surveyed and together represent approximately 5% of the total size measure for each TG by GEO. The survey population was then initially stratified into a stratum that is sampled with certainty, at most two take-some strata and a take-none stratum. The method for stratifying the population was based on the Lavallée-Hidiroglou algorithm (1988). It was modified by Ferland (2003) to take into account the expected rates of out-of-business units.
- e) Inflated the optimal sample size to compensate for non-responding or misclassified units on the survey frame. For more information, see Trépanier (2004).
- f) Created sample and population update procedures to efficiently treat birth, death and misclassified units each month.
- g) Used the Hidiroglou-Berthelot (1986) method in the statistical edit procedure to identify outlying responses that can be influential or non-representative. The identified responses are excluded from impacting imputed values.
- h) Developed an imputation system that makes better use of historical and / or administrative data to impute for all non-responding units, especially for new units and dying units.
- i) Utilised Statistics Canada's Generalised Estimation System (GES) to quickly calculate estimate totals of the surveyed portion and their variances for various domains of interest. Also, a post-GES process was developed to calculate estimates for the take-none portion to allow for the total (surveyed plus non-surveyed) estimate to be determined.
- j) Developed analysis tools to better identify influential units causing large trend estimates in TG and GEO domains and to effectively treat them, if necessary. See Matthews and Bérard (2002).

- k) Developed diagnostic tools (they are described in sections 2.2 and 3).

## 2.2. Where are we now in the redesign project?

The second fiscal year of the redesign (2000-01) was devoted to the development of the sampling plan. The strategy for the edit and imputation as well as the estimation methodology were drawn up in 2001-02 and later finalized in 2002-03. In the early part of 2003-04, testing of the new processing systems was performed as well as the implementation of a pre-contact survey for those units that were selected to be part of the first sample for the redesigned survey. The first sample for the redesigned survey was drawn in April 2003. Recently a parallel run of five months was undertaken for the reference months of December 2003 to April 2004.

In the parallel run, the former MWRTS and the redesigned MWRTS were both administered from start to finish with monthly estimates being generated from each design for the five reference months. The parallel run was beneficial in terms of verifying that all the systems for the redesigned survey were functioning as expected on live respondent data from the newly selected sample.

Moreover, the estimates for a particular month (level estimates) and those comparing the particular month to an earlier month (trend estimates) from the former survey were available as a guide for those estimates resulting from the redesigned survey. Differences in sampling design had to be taken into account. For instance, the sampling frames differed since the coverage of the frame for the redesigned survey is larger including many businesses without employees. Businesses without employees (considered non-employer businesses) were, for the most part, not part of the former sampling frame. Non-employer businesses were taken into account each month in the former retail survey through the use of an adjustment factor that was constant throughout the years. The redesigned survey also has a take-none portion where the estimates come from administrative sources, while the former survey did not have such a take-none portion. The samples for the former and redesigned surveys contained a large overlap of units, although this was not controlled at the time of sample selection.

The month of April 2004 was the last reference month of the five month parallel run. During the first four months of the parallel run, only the estimates resulting from the former survey were published but, beginning in reference month April 2004, the estimates resulting from the redesigned survey were published. For the

five months of the parallel run and the months following, a number of survey processes were implemented to ensure proper monitoring of the various survey steps in the redesigned survey. Three different types of 'diagnostic' tools were developed with the following objectives:

- 1) To assess the functionality of the modules in each survey step.
- 2) To produce monthly descriptive statistics that can be studied longitudinally to detect changes that ultimately affect the estimates.
- 3) To perform analysis of the level and trend estimates.

During the months of the parallel run, the first objective was the one that was the most critical as it was important to have all the modules functioning as intended. However, as the parallel run was coming to an end, the last two objectives moved to the forefront as it became increasingly important to monitor changes that were driven by the data (either reported or imputed). The focus of the next two sections will be to present the diagnostic tools that were developed to meet the last two objectives.

### 3. EXAMPLES OF DIAGNOSTIC TOOLS

The diagnostic tools which are associated with numerous survey steps of the redesigned MWRTS were tested before and during the five month parallel run. Since the parallel run, fine-tuning of the tools to meet the needs of all involved in the survey is being undertaken. The three survey steps that this paper will concentrate on are 1) sample update, 2) imputation, and 3) estimation.

#### 3.1 Sample Update

The MWRTS is a recurring monthly survey that produces both level estimates and trend estimates (change estimates) for two variables: monthly sales and inventories. To produce accurate level estimates, it is necessary to update the initial sample on a monthly basis to take into account births and deaths in the population as well as structure and activity changes among the recurring units. However, the monthly sample updates must have a minimal impact on the monthly change estimates and must take into account survey feedback to the survey frame in order to avoid introducing a bias in the estimates. A weakness of the former survey was its lack of monitoring at the sample update stage. Without efficient tools at this stage, it was not rare to have surprise discoveries at the estimation stage.

It should be noted that in April 2003, the sample for the redesigned survey was first selected and it was done independently of the sample that existed for the former survey. This period is referred to as original stratification. Furthermore, the sample for the redesigned survey remains virtually the same month after month with the exception of a sample of birth units that are added monthly.

Statistics Canada's Business Register (BR) provides the survey frame for the MWRTS. The sample updates are driven by this survey frame that is updated each month. The BR is a structured list of businesses engaged in the production of goods and services in Canada. It is a centrally maintained database containing detailed descriptions of most businesses operating within the country. The businesses on the BR are represented by a hierarchical statistical structure with four levels, with the statistical enterprise at the top, followed by the statistical company, the statistical establishment and the statistical location. An enterprise can be linked to one or more statistical companies, a statistical company can be linked to one or more statistical establishments, and a statistical establishment can be linked to one or more statistical locations.

The MWRTS sampling unit is called a *cluster* and it is a group of statistical establishments. A cluster of establishments is defined as all establishments belonging to a statistical enterprise that are in the same TG and GEO. The MWRTS is actually two separate surveys with two independent samples (one of establishments classified to the retail trade sector and the other of establishments classified to the wholesale trade sector). The size of the sample for the retail trade survey is around 12,000 clusters while it is approximately 8,000 clusters for the wholesale trade survey. Each month when the sample is updated to account for survey frame changes, the diagnostic tools that have been utilised during this step have proven to be very valuable.

Before proceeding into describing the various sample update diagnostic tools, it is important to mention that a user friendly system has been developed using SAS/AF to quickly access the diagnostics.

The following image (Image 1.) shows the menu screen of this system. From this menu, it is possible to quickly access various diagnostics from the particular survey (either retail or wholesale) for a particular reference period. For a number of the options that can be selected from the menu screen, it is possible to have the diagnostics in tabular or graphical form.

**Image 1: Sample Update Diagnostics Menu Screen**



One of the first diagnostic tools (labelled ‘Live Est’ in Image 1.) used each month following the execution of the sample update process in GSAM is the comparison of the number of live statistical establishments to the previous month. This is an important tool as it gives a heads-up on what to expect when data are received from the respondents for that reference month. This comparison is done for each industry and geographical region level. For instance, assuming all other factors being constant, if there has been a significant increase in the number of live establishments in a particular industry from one month to the next, an increase in the monthly sales estimate would not be considered a surprise. It should be noted that some month-to-month fluctuations result because of the way the BR makes its updates and these movements should be taken into account during any analysis.

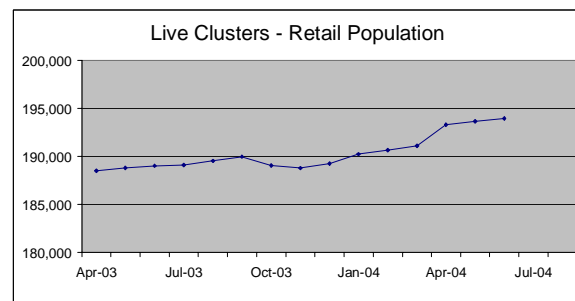
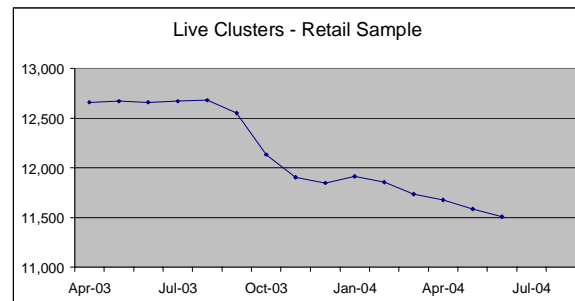
Another important tool (‘Clusters’) is one that calculates cluster counts. As mentioned earlier, groups of statistical establishments belonging to a statistical enterprise that are in the same TG and GEO are known as clusters. Clusters are the sampling unit and they are, for the most part, the level for which data are collected from the respondents. Reporting preferences and data availability by the survey respondent can result in collection entities being different than the cluster. The cluster counts for the population and sample are provided in terms of live clusters (containing at least one live statistical establishment) and in terms of total clusters (live and dead). Sampled clusters that die (become either out-of-business or out-of-scope) remain in the sample since they are representative of the population. However, they are important to operations staff since it will not be required, after a verification, to contact them for a response. Contrarily, new clusters (e.g. those containing just new statistical establishments) can enter the sample each month. It is vital that

operations staff know the number of units that will have to be contacted for the first time.

Chart 1 that follows is a graphical example of how the number of live clusters has changed from April 2003 to July 2004 in both the retail sample and population. As one can see, the sample of live clusters has generally been declining during this period while the population has shown a general increase. The larger drop in sampled clusters during the fall of 2003 (just prior to the 5 month parallel run) is due to a pre-contact survey that was administered to the newly selected sample for the redesigned MWRTS in order to access their status (ie. industry, geography, live/dead). Many deaths were discovered in the sample during the pre-contact survey. This result showed that the administrative source was not as likely as identifying the dead out-of-sample clusters as the pre-contact survey was in identifying the dead in-sample clusters. From February 2004 to July 2004, the decline in the number of live clusters in the sample has been due to the fact that there have been more dead clusters in the current sample than new live clusters entering the sample as births. The decline during this same period is not present among the live clusters in the population since deaths are not as readily known among the non-sampled units.

A number of other charts for other types of clusters (i.e. total, dead, etc.) are available through this diagnostic tool in graphical as well as in tabular forms.

**Chart 1: Number of Live Clusters - Retail**



A number of other diagnostic tools for the sample update involve the size measure of the clusters. As mentioned earlier, the MWRTS population was initially stratified into a stratum that is sampled with certainty, at most two take-some strata and a take-none stratum. The stratification is done by TG and GEO and by a size measure that is created using either independent survey data or administrative variables. Take-none strata serve to reduce respondent burden by excluding the smallest clusters from the surveyed population. In the redesigned MWRTS, these clusters together represent at most five percent of total sales for each TG by GEO domain. Instead of sending questionnaires to these units, their contribution will be estimated through the use of administrative data.

Among the various diagnostic tools available at sample update, there is one ('TN to TA') involving the size measure that identifies take-none clusters that could possibly be promoted to a take-all cluster. This tool identifies candidates among the take-none clusters by listing those whose current size measure exceeds the boundary separating large take-some clusters from the take-all clusters (as opposed to what was determined at original stratification). These candidates are researched in order to determine if they should become a take-all cluster or not. This diagnostic tool has been developed to quickly identify poorly stratified clusters in our sample and to minimize the effect of take-none clusters upon the total estimate. Remember that the total estimate consists of the estimate coming from the surveyed portion (take-all and take-some clusters) plus the non-surveyed portion. Through monthly sample updates, it is desired that the contribution of the non-surveyed portion to the total estimate does not significantly exceed five percent, especially in the months following the onset of estimates resulting from the redesigned survey. With the amount of changes to the MWRTS, it was decided to keep a tight control on take-none clusters since it was desired that the surveyed portion should be principally driving the monthly trend estimates.

Another useful tool ('Size Chg') identifies clusters, based on a pre-specified condition, having size strata changes. The MWRTS has a pre-specified condition of making clusters that consist of multiple industry groups and/or multiple geographic regions take-all clusters regardless of their size. For example, this can result when a sampled take-some cluster begins reporting for two or more TGs or GEOs because there has been a change in its structure (which results in this cluster being different from other clusters in its earlier stratum in terms of TG and/or GEO).

Whenever a cluster changes its stratum as above and a cluster is promoted to the take-all stratum, the effect on the estimates is monitored in order to determine if the resulting estimates mirror reality. The effect is even more closely monitored when any restratification is performed. For instance, it is planned in the redesigned MWRTS to do an annual mini-restratification based only on size measure changes to restratify the most extreme cases. In this type of restratification, the resulting listing from another diagnostic tool will indicate the magnitude of size strata changes as well as the reasons behind the changes. For more information on restratification, see Majkowski (2001).

A related tool ('Size Table') outputs cluster counts from the population for each stratum based on the strata that would result from using the size measures of two different periods. These tables can indicate when it would be a good idea to do a restratification based on size measure changes.

The final two sample update diagnostic tools ('Max Wgt' and 'Wgt Chg') to be presented in this paper involve the sampling weights for the clusters. At original stratification, maximum weights that took into account the type of industry were set for the large take-some (LTS) and small take-some (STS) strata. The maximums were created to minimize the occurrence of erratic estimates that can result if a sampled unit's response is unrepresentative of non-sampled units. The sampling weights are recalculated each month based on the changes to the population / sample for the current reference month, while GSAM attempts to get the sampling fraction as identical or as similar to the sampling fraction at original stratification. Since rebalancing of the frame is not performed each month, it is possible to get sampling weights somewhat different than at original stratification. For more information on GSAM and rebalancing, see Statistics Canada (1996). The TG by GEO by size strata that have larger than maximum weights are monitored at the TG by GEO cell level in order to determine if they still produce reliable estimates. The other tool identifies cells that have a large percentage change in the sampling weight when compared to the previous month. The benefit of this tool is to prevent large trend estimates resulting somewhat artificially if they have been simply caused by a difference in sampling weights from one month to the next. If the sampling weight ends up considerably higher than the maximum and the reliability of the estimates also decreases, then this also signals that some sort of restratification is necessary.

An example of the maximum weight diagnostic tool is shown in Table 1. It shows which TG x GEO cells have a stratum (either LTS or STS) with a sampling weight (wgt) that exceeds the specified maximum (Maxwgt). The weight is calculated each month with the updated population (N) and sample (n) counts in each stratum. As can be seen for all the cases in Table 1, the sampling weight has only slightly exceeded the maximum weight. These cases will be monitored each month.

**Table 1. Maximum Weights Exceeded – Wholesale (May 2004)**

TG	GEO	Stratum	N	n	Wgt	Maxwgt
030	35	LTS	151	15	10.07	10
030	35	STS	601	20	30.05	30
050	24	LTS	217	36	6.03	6
050	59	STS	182	12	15.17	15
062	24	LTS	81	8	10.13	10
062	24	STS	361	12	30.08	30
062	35	LTS	138	13	10.62	10
090	24	LTS	62	6	10.33	10
100	48	STS	250	8	31.25	30
130	24	LTS	223	43	5.19	5
160	24	STS	1590	63	25.24	25

### 3.2 Imputation

Imputation in the MWRTS is performed for those responses that are missing for reasons such as the respondent not being reached, the data not being provided when the respondent was contacted or a refusal to respond for a particular month. The monthly questionnaire nonresponse rate based on clusters in the redesigned MWRTS has been between ten and fifteen percent at the national level. The diagnostic tools for the imputation process have been beneficial in assessing the quality of the values derived by the imputation system to replace the missing values.

The strategy that is used by the imputation system to impute a missing value is dependent upon whether the cluster has been previously in the sample or if it is sampled for the first time (a birth cluster). For those non-birth clusters, historically reported responses as well as general trend techniques are used as a basis for the imputation. General trend imputation requires the identification of imputation groups, i.e., the identification of homogeneous groups of clusters at the TG, GEO and size level. Since historical data from the survey are non-existent for birth clusters, administrative data (i.e., data from the GST program) as well as a trend calculated from similar clusters that responded is used as the imputation method. A diagnostic tool that summarises which imputation method has been used is helpful each month in determining if the reliability of the estimates is influenced by the type of imputation method used.

This will become more useful once the total variance of the estimates includes a measure indicating what proportion of the variance is due to the imputation system. This is a future project for the MWRTS.

Table 2 is an example of a summary table showing the frequency of imputation methods that were used to impute a value for wholesale sales in March 2004 for 4 TGs. Note: Ratio imputation using previous month (PM), ratio imputation using previous year (PY), ratio imputation using next month (NM), regression using previous month and previous year (RM) and ratio using GST information (GST).

**Table 2. Imputation Method vs. TG (Wholesale Sales – March 2004)**

TG	Imputation Method					Total
	PM	PY	NM	RM	GST	
110	25	1	0	7	0	33
120	5	0	1	21	0	27
130	54	35	1	84	0	174
140	11	11	1	29	2	54

A second diagnostic summarises which imputation group was used. Following imputation, a summary of imputation groups used can provide some insight into modifying imputation methods or types of imputation groups. Often it can be due to the fact that the trends were derived from responding clusters whose responses are quite different from the non-responding clusters.

### 3.3 Estimation

Estimation is the survey step that assumes that input values (following imputation) are correct and complete before calculating the required estimates and their respective quality measures. In the redesigned MWRTS, estimates and quality measures are calculated for the survey portion through the use of GES (Estevao, 2004). For the non-surveyed portion, estimates derived from administrative data are calculated by a post-GES process. In this process, the total estimate (surveyed plus non-surveyed) is also calculated. The quality measures from the surveyed portion are transferred to the total estimate. This is done for the time being since it is planned to better estimate the error associated to the non-survey portion in the future.

For each reference month, a listing of clusters is produced containing the top contributors to the level estimate. The importance of this diagnostic is to verify that this list does not contain any surprises in terms of composition and to verify that the reported or imputed value for these influential clusters is accurate. Another diagnostic related to the preceding

one performs a comparison of top contributors from one month to the next. The appearance or absence of top contributors can result in a large monthly change estimate. Again, research is required to verify the accuracy of influential clusters.

Sometimes take-some clusters can be poorly stratified and can lead to inconsistent results. A diagnostic tool can identify the clusters that are not representative of other take-some clusters. In extreme cases, it is important to promote this cluster to the take-all stratum with a weight of one and to adjust the weights of the remaining take-some clusters accordingly.

In order to assess the impact of the various treatments performed on the clusters, a diagnostic tool has been developed to determine the proportion of the total estimate that is due to certain populations of interest. Some population of interests are the following.

- a) Birth clusters
- b) Rebirth clusters (clusters previously alive that died and have recently become alive)
- c) Imputed clusters
- d) Re-stratified clusters (has a stratum or weight change been significant?)
- e) Pre-specified clusters (the effect of making a cluster pre-specified)
- f) Clusters that are adjusted by a correction factor

A final diagnostic identifies the top contributors leading to a coefficient of variation (CV) for an estimate that is above a pre-determined target at the national, industry group or geographical region level. This diagnostic is related to the outlier detection strategy (Matthews and Bérard, 2002) following estimation where influential clusters are identified as contributing to an unacceptable trend estimate. In the determination of the sample, it was chosen such that the expected CV through the use of the size measure at original stratification would be less than the pre-determined target. It is planned for the future to incorporate this tool with the outlier detection strategy into one tool in order to identify the clusters in cells which have unacceptable realised CVs, to determine those causing unacceptable trend estimates, and to suggest treatments for them.

An example of the output produced by this final tool is shown in Table 3. The table shows the last part of this tool following the identification of units leading toward higher than target CVs and following their suggested treatments. Table 3 contains the effect of implementing the suggested treatment on the estimates and CVs for selected TG x GEO cells. It should be noted that the expected CV was calculated

at the time of initial stratification (April 2003) and is based on the size measure at that time.

**Table 3: Outlier Treatment (Wholesale Sales – March 2004)**

TG x GEO	Target CV (%)	Expected CV (%)	Before		After	
			Est. (M\$)	CV (%)	Est. (M\$)	CV (%)
000 x 00	1.20	0.67	38,592	0.73	38,284	0.67
000 x 10	2.50	1.17	223	2.41	217	1.37
000 x 11	2.50	1.93	49	11.84	43	1.83
110 x 00	3.50	2.60	1,181	4.24	1,170	3.77
120 x 00	3.50	2.94	1,054	3.51	1,047	3.18

As the Table 3 shows, there is one TG x GEO cell (110 x 00) where the CV after implementing the suggested treatment still results in one that is above the target (3.77 versus 3.50). This has resulted since the auxiliary information used in the outlier detection strategy does not clearly show that further treatments are necessary in this cell. If this particular cell continues having CVs above the target, it is a signal that either the sample in this cell needs to be augmented or renewed.

**4. HOW ARE THEY USED IN PRODUCTION?**

The tools that have been described in this paper (as well as others that were not described) have been created for the redesigned MWRTS to be easily accessible and to produce outputs that are useful for analysis purposes. For many of the outputs produced, their timing in the monthly production cycle is critical for them to be of value. For example, let us consider a cycle for one reference month and describe how the process currently works.

Upon the monthly execution of GSAM, a new updated sample is chosen. This sample will be similar to the previous month except for such changes as birth clusters entering the sample and existing clusters having different weights due to structural changes. The updating of the sample as well as the running of the diagnostic tools for sample update is performed at the start of the reference month. For example, for the reference month of January 2004, the first week of January was for the sample update to be run and for the sample update diagnostics to be produced. The diagnostics were utilised during the rest of January in preparation for the collection of data during the following month. The collection of January data was terminated at the end of February. During the first week of March, imputation and estimation were performed with the diagnostics for each one being run following their execution. If the diagnostics revealed a problem following imputation, there was time to make minor modifications and re-run imputation or to manually

implement a change prior to running estimation for the first time. Following the running of estimation, the estimation diagnostics were immediately produced. There was at least one more run of estimation prior to releasing the estimates and CVs during the third week of March.

For the sample update diagnostics a user-friendly menu-driven system built with SAS/AF was utilised. From the menus, it was possible to quickly choose which diagnostics were desired and to run diagnostics not only for the current month, but for any other past month desired. Currently, such a user-friendly system does not exist for the imputation or estimation diagnostic tools, but the desired diagnostics can be produced by running a set of SAS programs.

## 5. GENERAL CONCLUSIONS

During the parallel run of the redesigned MWRTS with the former MWRTS, the diagnostic tools that were developed were for the first time run during the normal survey cycle. It was probably not the ideal time to determine which diagnostic tools were the most vital since time was somewhat limited as the priority was still with the former MWRTS (at least for the first four months). Nevertheless, several ideas / lessons presented themselves.

The user-friendly system built for the sample update diagnostics was found to be a useful tool in quickly assessing the characteristics of the sample and the sample updates. A similar kind of system needs to be built for the other survey steps so that important diagnostics can be readily available in order to avoid slowing down the overall processing of the data.

For efficiency, there is a need to identify which diagnostics should be run each month. With feedback on the current diagnostics we will have a better idea of which ones are the most vital and which diagnostics must be made available on request. All diagnostic tools need to be fully documented in order to avoid misinterpretations.

Finally, the diagnostic tools that have been developed for the redesigned MWRTS have played an important role during the processing of survey results. They have minimized problematic situations and have explained the various findings prior to the calculation of the final estimates and the publication of the data. Without these tools, it would have been much more difficult to determine how the sample / population changed each month and how these changes have affected the resulting estimates.

## 6. ACKNOWLEDGEMENTS

The author would like to thank the co-authors, Steve Matthews and Sonia Demers for their contributions and helpful comments on this paper. Furthermore, Michel Ferland, Jean-Sébastien Provençal, Robert Philips and Sébastien Labelle-Blanchet also need to be acknowledged for their contributions to the project.

## 7. REFERENCES

Bérard, H. (2001). "The Redesign of the Monthly Wholesale and Retail Trade Survey of Statistics Canada", *Proceedings of the Survey Methods Section*, Statistical Society of Canada, pp.81-86.

Estevao, V.M. (2004). "GES V4.3 User Guide", Statistics Canada report.

Ferland, M. (2003), "Enhanced Lavallée-Hidiroglou Algorithm", Unpublished manuscript, Ottawa: Statistics Canada.

Hidiroglou, M.A. and Berthelot, J.-M. (1986), "Statistical Editing and Imputation for Periodic Business Surveys", *Survey Methodology*, Vol.12, pp.73-83.

Lavallée, P. and Hidiroglou, M.A. (1988), "On the Stratification of Skewed Populations", *Survey Methodology*, June 1988, Vol.14, No.1, pp.33-43.

Majkowski, M. (2001), "Maintaining Estimate Quality and Easing Response Burden in a Sub-Annual Business Survey", *Proceedings of the Survey Methods Research Section*, American Statistical Association [CD-ROM].

Matthews, S. and Bérard, H. (2002), "The Outlier Detection and Treatment Strategy for the Monthly Wholesale and Retail Trade Survey", *Proceedings of the Survey Methods Section*, Statistical Society of Canada [CD-ROM].

Statistics Canada (1996). Generalized Sampling System (GSAM) User Guide, Ottawa: Statistics Canada.

Trépanier, J. (2004), "The Redesign Canadian Monthly Wholesale and Retail Trade Survey: A Postmortem of the Implementation", *Proceedings of the Survey Methods Research Section*, American Statistical Association [CD-ROM].