Exploring the Usability of GPSed Records: A data typological approach*

Hiromi MORI†

Summary

Up until quite recently, location information on surveyed units, for example, of households, establishments and enterprises, has been collected as area information, such as tract codes and municipality codes, in which one-to-one correspondences between the unit and the location information are not provided, despite the fact that each unit has inherently unique information in terms of its location.

This paper first addresses the issue that ambiguity of data due to insufficiently obtained location information under questionnaire-based surveys gives rise to several constraints in their use. Latest developments in information technologies have opened up new possibilities for the application of GPS for statistical purposes. One can create GPSed records by assigning relevant GPS codes to respective survey results. Compared with non GPSed records, GPSed records appear to yield several benefits. Thus, the remainder of the paper highlights the potential function of GPS codes with respect to the possibilities of cross-sectional as well as longitudinal data fusion, which is expected to explore new frontiers in integrated data production by expanding the dimensions of existing records. The discussion in this paper also implies that GPS coordinates are one of the possible key variables applicable to the integration of data.

Keywords: GPS, data archive, data fusion, data integration

1. Introduction

In modern census, enumerating activities have been conducted at each census

---

* This paper is based on a presentation "Exploring Usability of GPSed Records- A data typological approach" made at the workshop "Statistical Innovation: Use of GPS and GSM data and integration" organized by Central Bureau of Statistics Netherlands on September 6, 2010 in Heerlen. An earlier version of this paper was already published in March 2011 on *Statistics* (The Japan Society of Economic Statistics), No.100.

† Faculty of Economics, Hosei University
4342 Aihara, Machida-shi, Tokyo, JAPAN 194-0298

tract, which exclusively covers the whole scope of national territory. Census tracts were introduced with the intent of avoiding counting failures as well as multiple counting. Although some surveys inquire about respondents' addresses, in most cases, location information of surveyed units is given either by region or tract code. The tract code has been used as the minimal unit to represent the location information of surveyed units within the tract. In other words, surveyed units in a particular tract have shared identical code numbers to represent their locations.

Modern information technology has provided a substantial breakthrough in obtaining the location information of surveyed units. Due to advanced information technology, together with the wide-spread use of reasonably-priced handheld PCs, the Global Positioning System (GPS), originally introduced as a military invention, is now widely applied in various fields as a civilian technology.

The aims of this paper are twofold: first to document a set of problems caused by the insufficient collection of the location information of surveyed units, and second to draw a sketch of the potential uses of GPSed records with regard to the typology of data.

2. Statistical surveys and the dual nature of surveyed records

In conducting surveys, information is collected from surveyed units such as persons, households, establishments, enterprises and so on, through the use of questionnaires. The information obtained concerning surveyed units is usually arrayed as a record format, which, however, has a dual nature.

It is obvious that the obtained data, i.e. the various attributes, activities and results, are ascribed to each surveyed unit. That is, individual records have been regarded as statistical copies of the surveyed unit. Another aspect is less obvious compared with the first one. The surveyed information belongs to or relates to the units that are located at a particular geographical point, i.e. a dwelling unit or site where business activities are carried out. Put differently, a set of informational data offered by surveyed units are related to some particular geographical point. One may term the former "unit information" and the latter "spot information."

Spot information obtained from observations in a single survey is less obvious than unit information, because spot information refers not to the unit itself but to its locational existence. Repeated observations, however, may more clearly address the dual nature of the records. When the same unit has been repeatedly observed in a series of surveys or censuses, the obtained records may reflect longitudinal change in the relevant unit. When the same spot has been observed in repeated surveys, it will document the kind of activities of one fixed point at different moments.

As these two aspects which the surveyed records inherently possess in a latent manner are substantially dynamic in nature, they may split off in cases when units change their locations over a period of time. Although the majority of the surveyed units continue to stay at the same spots, the replacement of units may possibly take place in surveys conducted at certain intervals. Different units may be observed in ensuing surveys at the same spot due to the replacement of units, i.e. by a former unit moving out followed by a substitute moving in. The observed spots in the previous survey can disappear, whether or not the dwelling units are existent, in cases when no succeeding tenants accommodate that dwelling unit. It may also be possible that new entrants are surveyed at new spots. Families can be occupants either of newly constructed or formerly being unsettled dwelling units, while establishments and companies can launch their business activities either at newly developed industrial sites or ones that were formerly unoccupied when the previous survey was conducted.

Statistics has long been regarded as a science dealing primarily with massive phenomena. In traditional statistics, therefore, surveyed units used to be regarded simply as elements that mold a population or subpopulation. It was only in the latter half of the 20th century that statisticians began to shed light on individual survey records.

Due to these traditional statistical ideas, together with several technological constraints, statistics remained tolerant of the insufficient use of the location information inherent in survey records. Although surveyed units such as households, establishments and enterprises mostly have definite location information regarding their existence, survey records documented them not at their particular points, but only as one of the component units of the tract. Instead of specified location codes inherent to respective surveyed units, a tract code number was given to all surveyed units that belonged to a particular tract. Each unit's location information was collected not as a geographical point, but as small area. Because insufficient location information was obtained, statistics had to put up with "diluted" information in terms of the location of units that resulted in a number of constraints on its use. Figure 1 illustrates examples of traditional household and establishment/enterprise record layout forms.

Household survey record

	date of survey		**location codes**					survey items			
survey identification code	year	date	**prefectural code**	**city code**	**survey tract code**	family sample number	seq. number of ind. in family	item 1	item 2	item 3	・・・

Enterprise/establishment record

	date of survey		**location codes**			survey items									
survey identification code	year	date	**prefectural code**	**city code**	**survey tract code**	name	ZIP code	address	startup date	capital size	number of employees	item 1	item 2	item 3	・・・

Figure 1. Examples of record layout forms

3. Information constraints of tract-based records

(1) Problems caused by border rezoning

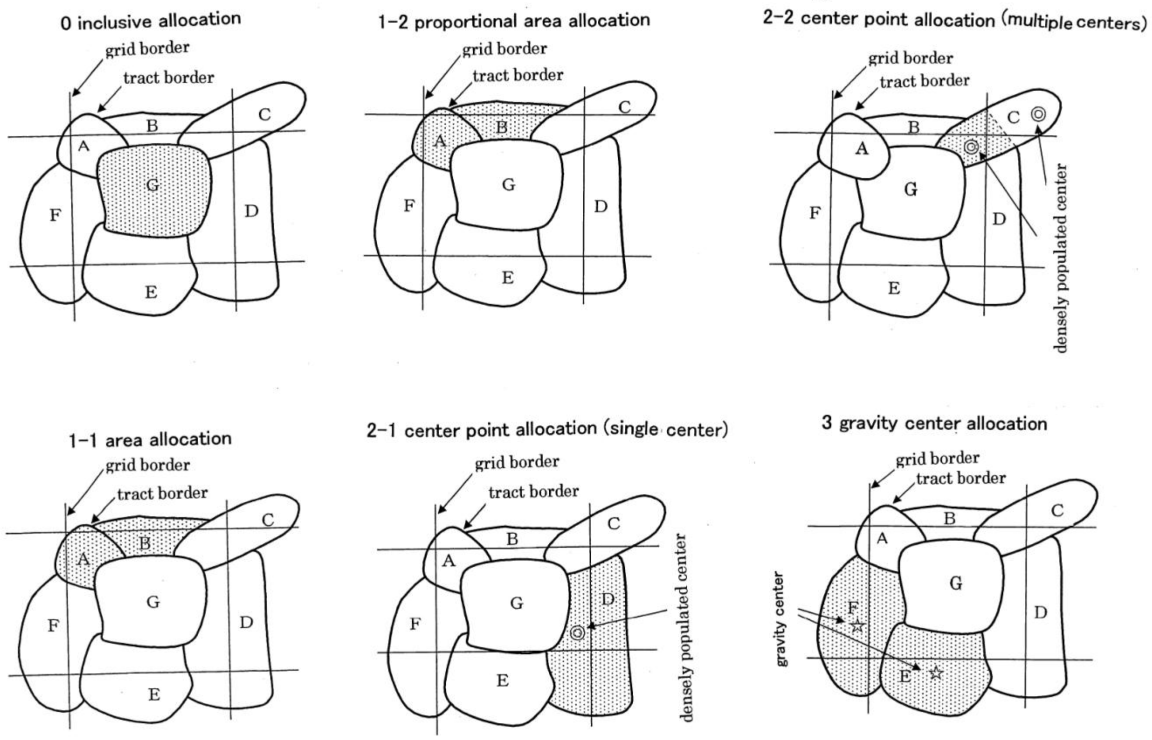
While Japan had more than 12,000 cities, towns and villages in the 1950s, the number had diminished drastically to about 2,200 by the year 2005. The annexation and reorganization of municipalities are real threats to statistical comparability, since they require enormous amounts of clerical work to adjust historical statistics to the newly arrayed boundaries. The rezoning of boundaries renders time series regional data less consistent.

Census tracts are not totally immune from boundary rezoning. The completion of new roads and railways and the development of new residential areas make existing tract maps obsolete. Some tracts have been partitioned and then annexed to several neighboring tracts, while several others have been totally reorganized. Such tract rezoning also disturbs the comparability of small area time series data. The Basic Unit Block (BUB) was introduced in Japan in 1990 as the minimal survey tract area of a more or less durable nature. Although these are expected to be more robust than census tracts, they still are not always free from restructuring.

(2) Allocation of surveyed units in tracts

Grid Square Statistics were introduced in Japan based on the 1970 census data. Under this system, the whole national territory is divided into rectangles of about one square kilometer and 500 square meters by longitudinal and latitudinal lines. These grids are called “basic grid squares” and “half grid squares,” respectively.

Since the geodetic line partitions areas mechanically into a set of uniform grids, they can be independent of any municipality rezoning and of tract reorganization. As case 0 in Figure 2 illustrates, for tracts that are totally included in a particular grid, the whole of their elements are properly allocated to that grid. In the case where the grid borders cross the tracts, however, tract elements, i.e. the surveyed unit records, should be processed in such a way as to cope with the problems of how to allocate them among grids in an appropriate manner. In all remaining cases, surveyed units are allocated more or less by approximation (case 1-1) or by calculation (cases 1-2, 2-1, 2-2 and 3). In either case from 1-1 through 3, an ambiguity occurs in converting tract-based data into grid-based data.



Source: http://www.stat.go.jp/data/mesh/pdf/gaiyo2.pdf (pp.24, 26 and 28)

Figure 2. Allocation of tract units among grid squares

(3) Inadaptability of data for buffering analysis

Buffering analysis is now widely used to identify buffered polygon areas with a fixed distance surrounding specified input features, which can be polygons, lines or points. Since buffer polygon borderlines do not necessarily coincide with those of tracts, borders usually intercross. Similar to the grid estimates, estimates for buffered polygons, therefore, are usually subject to the ambiguity caused by inconsistent borders. Buffered circles and polygons usually have indented fringes due to the discordance of bordering.

All these difficulties, yielded in the process of allocating surveyed units in tracts to relevant areas, derive from insufficiently obtained location information in surveys.

4. Obtaining GPS coordinates

Developments in information technologies have opened up a new scope in obtaining location information from each surveyed unit. Similar to the Internet, GPS was originally invented and has been utilized primarily for military purposes. Thanks to improvements in the accuracy of digital map software, together with widespread use of information terminals furnished with various GIS softwares, GPS now enjoys a wider acceptance in daily life as necessary information infrastructure.

Official statistics, however, are relative latecomers in applying GPS for their practices.

Directly obtaining GPS coordinates through mobile terminals and indirect access to them either by means of address-GPS converting software or by applying appropriate calculation methodologies served as a powerful driving force for statistics to explore the wider dimensions of the applicability of coordinates, not only for the use of data but also for the production of data of improved quality. GPS coordinates collected by field workers in address canvassing for the 2010 U.S. Population Census are used to pinpoint the residences of non-responders, and thus to improve the response rate. Statistics Poland is planning to collect GPS coordinates in the 2011 Census.

Besides such applications of GPS coordinates in the survey process, they are expected to provide a wider dimension of inputs to statistical practices. As one of the major aims of this paper is to address the characteristics of individual records with GPS coordinates (hereinafter termed GPSed records), it would be convenient to provide here a rough image of GPSed records. The diagrams in Figure 3 document images of a data format for GPSed records.

Household survey record

	date of survey		location codes						survey items			
survey identification code	year	date	prefectural code	city code	**GPS coodinate X**	**GPS coodinate Y**	family sample number	seq. number of ind. in family	item 1	item 2	item 3	･･･

Enterprise/establishment record

	date of survey		location codes				survey items						
survey identification code	year	date	prefectural code	city code	**GPS coodinate X**	**GPS coodinate Y**	name	ZIP code	address	startup date	capital size	number of employees	･･･

Figure 3. Examples of GPSed records

Unlike tract-coded records, GPSed records provide definite location information of surveyed units. As stated above, ambiguity in the use of data has sprung substantially from area-based locating. GPS coordinates are more appropriate variables than tract codes in terms of identifying the geographical points of surveyed units' existence. Once GPS coordinates are tacked to individual records by some measure or other, it becomes possible to allocate surveyed units not by estimation but by direct assorting of surveyed units according to the coordinate information. Units such as families, establishments and enterprises will be surveyed intrinsically at the very point of their existence. It was not until the obtaining of coordinate information that statistics became able to employ location information on an extensive scale.

GPS coordinates tacked to each record as one of the unit's basic attributes will enable to liquidate the ambiguity described above. By doing so, all archived records will be able to cope with any form of zoning. GPSed time series records can enjoy longitudinal comparability in full scale. Furthermore, they are qualified to compile statistics that can meet any buffered zones.

Besides these advantages, GPSed records appear to have additional attributes. The following paragraphs will discuss potential uses of GPSed datasets with regard to the typology of data.

5. GPSed records by type of datasets

Datasets can be classified into several subcategories by kinds of surveyed units and forms of datasets. Additional variables that account for the datasets will also be introduced to characterize the specific nature and usability of GPSed datasets.

A single census or survey result provides a snapshot of the surveyed units at a particular date that forms a single cross-sectional dataset. A series of censuses or surveys conducted repeatedly during the sequence of time will give repeated snapshots. These snapshots usually comprise repeated cross-sectional datasets. Leaving aside censuses, a series of survey results do not necessarily cover the same surveyed unit. Repeated cross-sectional datasets, therefore, do not portray snapshot observations of the same set of surveyed units. When the same units are surveyed repeatedly in a series of surveys, one can compile longitudinal datasets that form a matrix of N surveyed units and T periods for each surveyed variable. However, the number of surveyed units in each snapshot is not always the same in the longitudinal dataset because of the attrition of the surveyed samples. Including unbalanced datasets with an unequal number of surveyed units in each snapshot, the author simply terms such datasets here as longitudinal.

As for the nature of surveyed units, we will focus our discussion on the GPSed records of surveyed units with a rather stable nature in terms of their geographical locations. Thus, locations, i.e. dwelling units usually inhabited by families and sites where establishments/enterprises perform their economic activities, are currently our major concerns in discussing GPSed records. Individual records loaded with GPS coordinates involve in themselves a potential moment to breakaway the dual nature that seems to be inseparably integrated in the surveyed records. This separation will turn out to be pronounced in the repeated snapshot datasets such as repeated cross-sectional and longitudinal datasets.

Table 1 Business/household datasets by type

surveyed units	observation unit	single snapshot	repeated snapshots	
		cross-sectional	repeated cross-sectional	longitudinal
business (enterprise/ establishment)	unit	(A)	(C)	(E)
	site			
household	unit	(B)	(D)	(F)
	dwelling			

6. Possible uses of GPSed records by type of datasets

Categories of GPSed datasets (A) through (F) in Table 1 appear to have particular attributes regarding each surveyed unit and its location information, which govern the scopes and dimensions of their usability.

(A) Cross-sectional GPSed business datasets

As figures 1 and 4 have documented, a pair of GPS coordinates (x, y) corresponds to each surveyed record, while surveyed units with a traditional record format share the same geo-codes, such as tract and other area codes. In the latter case, whole units that fall within a respective area should carry an identical location code number, such as a tract code. GPSed records are distinguished from non-GPSed ones, among others, by a one-to-one correspondence of surveyed record with its location code. Since GPS coordinates provide an individual record with accurate pinpoint information in terms of each unit's location, GPSed records can be free from ambiguity in allocating units into respective regional areas that non-GPSed records were unable to do.

Allocating units in bordering areas to pertinent areas has been an extremely labor-intensive exercise in compiling grid square statistics. As cross-sectional GPSed datasets can cope with any regional zoning, it may be possible to complete it automatically with the help of coordinate information. It is quite reasonable that the Japanese Statistics Bureau converts address data to GPS coordinates in compiling Grid Square Statistics from the Establishment and Enterprise Census data. They can also handle any claims in elaborating polygons required in various buffering analyses.

Cross-sectional GPSed business datasets may be applicable, for example, to the following analyses. Firstly, they can provide effective datasets for analysis of various aspects of industrial clusters. The territorial location of clusters, their economic size and density by region and industry are of major concerns among geographers.

The U.S. Census Bureau was so quick in assessing damages caused by the hurricanes Katrina, Rita and Wilma with GPSed establishment records (Jarmin S.Ron and Miranda J., 2009), where GPS coordinates had not been obtained neither by direct positioning the sites by field surveyors nor through address matching , but through calculation measures they loaded individual establishment records with coordinates. This case study offers one smart example demonstrating the potential usability of GPSed datasets, for example, in the field of disaster damage prevention. Central and local governments of most countries have already furnished various hazard maps. One may easily assess the extent of damage by overlaying GPSed records on hazard maps using coordinates as linking keys.

It is worth noting that GPSed cross-sectional records also have an advantage in enlarging the information potential of data by means of expanding dimensions through data fusion. Among individual records from multiple sources such as censuses, sets of heterogeneous surveys and administrative records, there may exist some which carry identical coordinate information. However, such cross-sectional record linkages are "pseudo," because it is not necessarily the relevant business units that were combined with each other as unified records in extended dimensions. The latest developments in statistics have shed light on data fusion as one of the possible expansions of information potential. Records with a multiplied number of variables generated by the coordinate-based cross-sectional data fusion among heterogeneous business records may allow intensive analyses that a single set of records could never hope to achieve.

(B) Cross-sectional GPSed household datasets

Unlike tract coded records, which share an identical polygon code number among surveyed units, each household record in GPSed datasets usually has a unique location code relative to the coordinate information of the dwelling unit. Although multiple-floor apartment houses may possibly be codified by one and the same pair of coordinates, coordinates may still retain their validity as location indicator. GPS coordinates are also expected to expand their dimensions, for example, by introducing an additional variable that denotes floor information.

GPSed household datasets are more informative than tract coded ones in analytical usability, because they are qualified to accommodate themselves to a wide spectrum of regional zoning. One can estimate or assess the number of casualties from natural disasters such as floods and earthquakes by overlaying the GPSed records upon hazard maps. Statistical assessments of governmental services may also be possible by scoring accessibility to public facilities. GPSed household datasets capable of meeting any buffering analyses are also attractive to businesses in mining potential local markets by calculating the size, compositions, density and income distribution of subpopulations in relevant buffering areas.

Expanding the potentials of existing data by data fusing records is also valid for household records. Despite the pseudo manner of data linkage, the compiled datasets with multiplied dimensions of variables will enable intensive analyses that may bring about new findings.

(C) Repeated cross-sectional GPSed business datasets

Since coordinates are distinct in indicating the location of units, one can obtain results not by estimation but by the direct counting of units through a vector algorism applicable to any levels of polygons. GPSed records can display their advantages over other location codes, among others, in time series regional comparisons. Once individual records are archived with appropriate coordinates, they will become able to release the data from every constraint in time series comparisons that was formerly caused by restructured borders. Allocating units to each pertinent polygon by the help of coordinate information will make possible prospective as well as retrospective regional comparisons.

Repeated cross-sectional GPSed business datasets obtained by a series of surveys will offer users a periodical chain of snapshots on the activities of business units and behaviors. They can be applied to the analysis, for example, of the dynamism of an industrial cluster. With these types of datasets one can draw a series of pictures that illustrate the trend of diffusion or contraction of industrial clusters and can analyze business demographic events such as the entry or exit of units to or from the cluster.

One of the characteristic features of the repeated cross-sectional GPSed datasets is the possibility of longitudinal expansion of data dimensions. When we focus our interest on the location information of surveyed units given by the coordinates of sites where establishments or companies currently perform their activities, a new type of dataset, i.e. a pseudo panel dataset of establishments or companies will be compiled by fusing records by means of coordinates. The dataset is pseudo in the sense that establishments or companies that perform their business activities at the respective sites are not necessarily identical units. Business being performed at a particular site may alter by the exits of units followed by substitute entries during the period of time in question. However, as it is expected that an overwhelming majority of business units continue to carry out their activities at the same sites they have occupied in the past, we regard the compiled datasets as a panel in the broader sense.

When one changes ones viewpoint to the location where each unit was actually surveyed, however, both panel datasets are "genuine" in nature. Put differently, GPS coordinates are qualified to work as effective key variables to generate panel datasets out of unpaneled repeated cross-sectional datasets. Thus, panel-based analyses would be applicable to these types of datasets.

(D) Repeated cross-sectional GPSed household datasets

Repeated cross-sectional GPSed household datasets give a chain of snapshots focused on the activities and behaviors of families over the course of time. Thanks to the coordinates, the datasets can support any restructuring of regional zones. One can analyze various dynamic aspects of population and families by each region using this type of dataset. Comparison of the ageing tempo of populations by region, for example, is of importance for policymakers who are keen on reallocating budgets.

When one regards repeated cross-sectional GPSed household datasets from the viewpoint of GPS coordinates, individual household records are reorganized into pseudo panel datasets. Similar to the business datasets, those compiled from repeated cross-sectional GPSed household datasets are still pseudo in terms of longitudinal attributes, because coordinates are linked not to respective families, but only to the dwelling units. Even in cases where household records carry unchanged coordinates in repeated cross-sectional datasets, there may possibly be replacements of families in dwelling units caused by the moving out of a family followed by moving in of another family. It is well expected, however, that in the majority of cases families continue to reside at the same dwelling units. Unless panel datasets in the true sense are available for households, pseudo panel datasets compiled by means of record linkage using GPS coordinates as matching keys would be applicable as one of the feasible options of a secondary approach to family demographic events analyses.

(E) Longitudinal GPSed business datasets

By the turn of the 21st century, business statistics in most countries had already become equipped with business registers that now serve as fundamental survey infrastructure as well as a particular machine to produce relevant statistics. Business registers in many countries have already stepped up to the second generation phase as databases with a longitudinal dimension in order to be able to meet the analytical needs of business demography. A business register, as the core segment of a relational database, forms a backbone for the integration of a wide spectrum of business statistical records both in cross-sectional and longitudinal dimensions. A systematic coding of the ID numbers of business units is a prerequisite for the effective functioning of the database. Longitudinal records in themselves contain elements of business demography, such as launching a business (entry), survival, dormancy (suspension) and quitting (exit).

GPSed longitudinal datasets are far more informative than non GPSed ones. Longitudinal records armed with GPS coordinates are qualified to objectify the dual aspect, i.e. unit and site information which the individual records have carried latently. When one focuses upon surveyed units, unchanged coordinates indicate their survival, while the changed ones suggest the redeployment of the unit. If one switches the viewpoint to sites, records illustrate the activities of the units operated at the particular site specified by the coordinates. Put differently, it will establish the kinds of functions or potentials of the respective sites.

GPSed longitudinal business datasets can identify the following events. When one focuses on the business unit in the dataset, its coordinates provide information regarding the unit's relocations in the course of time. Since the unit is identified by the competent ID number, one can easily distinguish redeployment from quitting.

GPS coordinates are more advantageous than descriptive address information in terms of data processing in identifying the redeployments of units. Addresses tend to be mistyped, while coordinates can maintain consistency even in cases when addresses are amended by occasional address recording.

Business units go through a set of demographic events throughout the period of their activities. When one focuses on the coordinates, surveyed unit records being identifiable by unit code number may denote the demographic events of the business unit, such as survivals, entries, exits which come about at a particular site. Thanks to the unit ID number, it is possible to distinguish new entries from the moving in of existing units due to redeployment and also exits from the moving out of units. It is expected that GPSed records can partly substitute for the profiling work of business units, which is actually quite labor-intensive clerical work, through automatic data processing.

By controlling site information, GPSed longitudinal business datasets would be applicable to establish, for example, the business unit redeployment ratio by size and industry and compare the ratios between single and multiple establishment businesses or grouped or single enterprises.

(F) Longitudinal GPSed household datasets

Building longitudinal household databases may currently remain a far-reaching project issue for most countries. However, Nordic countries have already switched over their statistical systems to register-based ones. Central Bureau of Statistics (CBS) of the Netherlands has constructed a modern version of the System of Social and Demographic Statistics (SSDS) as the Social Statistical Database (SSD), which is realized as a relational database with population register at its core segment and integrates many other household files as satellites.

As business registers have evolved, as a matter of course, from the first generation of the business frame that only reflected a static aspect of the business population to ones with longitudinal attributes, household registers will likely follow similar steps in the future. In this sense, the current status of statistical practices regarding household registers may be rather premature for the following discussion on the potential usability of GPSed longitudinal datasets.

Longitudinal household datasets can be compiled through matching records by family ID number. If no ID number is available, householders' names will substitute for the ID. Similar to the longitudinal business records, household records carry a dual implication. The record tells a story about the units themselves, i.e. families or individuals who share the dwelling unit on one side, and provides information on the functioning of respective dwelling units in terms of habitation on the other.

If we direct our concerns to units, i.e. families or individuals, a changed set of coordinates will trace the family or personal history of residential moves. This type of dataset is expected to provide relevant materials for analyzing the geographical residential moves of families or individuals in each stage of a family's or an individual's life cycle.

Longitudinal household datasets focused on dwelling units can draw another picture of the habitation behavior of residents. Household records reported from residents with unchanged coordinates may give either the same family ID number / the name of householder or differed ones in a series of snapshots. By overlaying family ID number or the name of householder on respective coordinates, one can compile a dataset that helps to shed light on the occupancy status of dwelling units. Unchanged ID numbers suggest that the same families or individuals continue to reside at the same dwelling units, while changed ID numbers indicate replacement of families or individuals. The coordinates that became extinct in the GPSed longitudinal datasets compiled of household-based survey results may indicate a vacancy or a halt of operation as residential dwelling units, while newly emerged coordinates suggest new engagements as residences. The datasets will be applicable to the identification of residential mobility, for example, by region and tenure.

7. Concluding remarks

Official statistics, which have collected information from surveyed units primarily to compile statistical tables, have experienced several historic turnabouts during the second half of the 20th century. Instead of macro-based datasets, the component of which are substantially aggregate statistics, users increasingly direct their concerns toward disaggregate data in the belief that the latter could portray novel images on population that aggregate data approaches were unable to attain.

Transition of the system of statistics from that made up substantially of stand-alone surveys to micro-based integration of surveyed and administrative records is another remarkable development. Collected information, which was formerly of temporary value for tabulating purposes, is more and more regarded as a sort of information asset of a durable nature that can meet long-standing and varied uses.

It is quite reasonable that contemporary needs for statistics require the archiving of obtained data which can withstand long term comparability and enable cross-sectional as well as longitudinal expansions of dimensions of archived records. The focus on GPS coordinates themselves in this paper derives from the anticipation that the loading of surveyed records with coordinates might be one of the possible options in designing data archives.

The use of GPS coordinates in official statistics is one of the recent remarkable developments in world statistics. Due mainly to technological and partly to social constraints, it was not until quite recently that statisticians began to turn their attention to the extensive use of location coordinates for statistical purposes.

Although questionnaire-based surveys are qualified potentially to collect information that belongs or relates to each surveyed unit, traditional surveys have offered location information in an aggregate manner as tract codes. Because of the ambiguity of available location information that arises from tract-coded records, survey results were subject to several constraints in use, especially in time series regional analyses.

This paper has given evidence for the following issues. First, besides information on the surveyed units, individual records also carry information about the locations of units in latent manner. Second, due to the aggregate nature of coding in terms of the location of units, transferred unit records from returned questionnaires, which are insufficient in giving the location information of units, have placed restrictions on their use. Third, as a typological approach has evidenced, repeated cross-sectional and, among others, longitudinal datasets can highlight location information given by the coordinates with special implications that may represent a sort of potential inherent in the respective spots. Finally, corresponding to each type of datasets, we have explored some new possibilities for GPSed records' usability.

As addressed by U.S. practices, GPS coordinates are also promising in producing better quality data. Dutch attempts may explore new potential uses in obtaining statistics otherwise almost unobtainable by conventional methods. Further seminal elaborations in the application of GPS coordinates in statistics may enrich the discussions being put forward in this paper. It is expected that new types of datasets armed with GPS coordinates will explore new frontiers in the field of statistics.

REFERENCE

Jarmin S.Ron and Miranda J., (2009) "The Impact of Hurricanes Katrina, Rita and Wilma on Business Establishments: A GIS Approach" *Journal of Business Valuation and Economic Loss Analysis*. Vol.4

Acknowledgement

This research has benefited from financial support through the Grant-in-Aid: "International Comparative Studies on Archiving System of Official Statistical Data" (#22330070) of Japan Society for the Promotion of Science.