TOPIC 12
GIS, LAND INFORMATION
SYSTEMS AND DATABASES

DEVELOPMENT OF A DISAGGREGATE LAND-USE ACTIVITY DATABASE

MARJALDI LOEIS
Transport Research Centre, RMIT University
PO Box 598 Market Street
Melbourne VIC 8007, AUSTRALIA

PARIS BRUNTON
Transport Research Centre, RMIT University
PO Box 598 Market Street
Melbourne VIC 8007, AUSTRALIA

ANTHONY J. RICHARDSON
Transport Research Centre, RMIT University
PO Box 598 Market Street
Melbourne VIC 8007, AUSTRALIA

Abstract

This paper describes the development of a land-use activity database for the city of Melbourne. The database contains information on employment in a range of industries (designated by ASIC division) at the geographic level of the census collectors district. Particular features of this database include spatially disaggregated data, the potential for regular updating, and incorporation into a geographic information system (GIS).

INTRODUCTION

Land-use information provides vital input into traditional as well as innovative transport planning models. It is used as the foundation for zonal trip generation and trip attraction calculations. It is again used heavily for trip distribution, and, to some degree, even for mode choice analysis.

What the transport planner essentially needs to estimate from a land-use information source is an indication of how many and what kinds of trips are generated and attracted by an area. The actual land-use designation of the area is useless unless that information can be utilised to determine some measures of socio-economic activity in the area, which in turn can be transformed into trip generation and attraction levels. Various indicators can be used to describe activity. ITE (1987) suggests such measures as number of employees, gross floor area, number of seats (eg in a restaurant), number of dwelling units (in a residential area) and others. The choice of indicator to be used depends on its suitability as well as data availability. In this work, the size of employment in the unit area was chosen because it is commonly available in different statistics. In addition, not only is it directly related to work trips, but it is also indicative of the size of activities taking place at a business site which in turn can be translated into the potential attraction (and generation) of the site for trips other than work trips.

The objective of the work reported in this paper is to develop a land-use database for Melbourne that contains area employment as a measure of activity level. In this paper, the requirements of the final product are first introduced, followed by the reasoning behind the choice of data sources and a description of those data sources. It then describes the procedures used for transforming the data into an employment database, and shows some results of this transformation. It concludes with an assessment on what has been achieved and how it can be improved.

THE LAND-USE ACTIVITY DATABASE

The development of this database was done as an integral part of the Victorian Integrated Travel, Activities & Land-Use (VITAL) toolkit project (Richardson and Ampt 1993). VITAL is a project that was initiated to provide a better information and analytical base to assist transport and land-use decision-making in the state of Victoria, Australia. The VITAL toolkit is an amalgamation of data and modelling processes that are systematically integrated within a framework of a geographic information system (GIS) package. This GIS framework enables consistent spatial definitions and allows seamless transfer of data between all the components of the toolkit. The land-use activity database is one of the components of this VITAL toolkit, and is therefore also governed by the GIS framework.

To further enhance data compatibility between the different components of the toolkit, land-use designations, particularly commercial land-uses, are standardised. For this purpose, the Australian Standard Industrial Classification (ASIC) system was chosen because of its widespread utilisation. Subsequent work will use the Australian and New Zealand Standard Industrial Classification (ANZSIC) as it has come to replace ASIC, but for the work reported in this paper, the older ASIC system has been utilised.

To enable maximum flexibility in the practical use of the data, it is necessary that this land-use activity database be as finely-grained as possible. In other words, the data is to be spatially disaggregated to the fullest extent possible so that it can be used most effectively. Furthermore, because land-uses constantly change, this database should itself be easily updated. To enable this, only regularly updated sources available in a practical electronic format can be considered for its input.

To recapitulate, there are five main criteria required by the final product, namely:

(i) It contains employment level as a measure of activity;

(ii) It contains an indication of the type of land-use using the ASIC system;

(iii) It is geographically oriented and can be integrated into the GIS framework of VITAL;

(iv) It is spatially disaggregated to a high degree; and

(v) It is practically updatable.

DATA SOURCES

To satisfy all the above requirements, non-traditional sources of data had to be used. For reasons that will be explained later, the main data sources that were decided upon were the Australian Bureau of Statistics' Business Register, the Telecom Business Finder™ (TBF) database and the Journey to Work (JTW) data from the 1991 Census of Population and Housing.

The Business Register

The Business Register is a database of businesses that employ wage and salary earners, and of agricultural operations that have an estimated value in excess of $5,000 regardless of whether or not they employ wage or salary earners (ABS 1992). In this database, the available data are the count of businesses classified according to their industry (ASIC code, all the way down to the lowest 'Class' level), location (by Statistical Local Area/SLA, or larger units of statistical area) and employment size (in seven standard employment size ranges). Also provided is the total number of employment units by industry code and location, except for agricultural units and for locations where the total number of business locations for a particular industry is five or less. One important drawback of this data source is that it does not include businesses (other than agricultural operations) that do not employ wage or salary earners. This automatically excludes from the data the small family or owner-operator type of businesses. In addition, it has been indicated that businesses that are no longer operating are often still recorded in the register. As a result, the recorded number of businesses and employment may be somewhat different (lower or higher) than the actual number in existence.

Of the five criteria to be met by the land-use activity database, the Business Register satisfies four. Firstly, it contains an activity measure in the form of employment size by geographic area. Next, it uses ASIC industry designations which fulfil the land-use type requirement and SLA divisions which allow geographical coding. Lastly, it is continually updated (new releases occurring about twice a year) and is provided in an electronic format. Its only problem is in the relatively large size of its smallest geographic unit (SLA). The SLA varies in size from 2 sq km in the central area of the city to 889 sq km in the outskirts. Although such zone sizes are quite acceptable for some transport planning purposes, the flexibility in the usage of the database may be highly limited.

The Telecom Business Finder™

The Telecom Business Finder (TBF) is an electronic version of the Yellow Pages directories. It is provided in the form of a CD-ROM containing information from all regional issues of the Yellow Pages in Australia (Telecom Australia 1992). As with the Yellow Pages, it includes descriptive and geographic information of individual businesses, including the type of business and the street address (in most cases). An important feature of the TBF is that its data can be easily exported into different formats, including text files that can be processed by many computer applications. The data can therefore be read and manipulated by other database management software, can be related to other databases, and because it contains individual street addresses, has the potential for being transformed into a highly disaggregated geographical database. Moreover, since it is updated annually like the printed version of the Yellow Pages, it also satisfies the 'updatable' criterion of the land-use database.

The TBF records follow the same business classification system as the Yellow Pages. In addition, matching ASIC codes are also provided for each record, therefore satisfying another criterion of the land-use activity database. Unfortunately, the assigned ASIC codes are based entirely on the Yellow Pages business classification under which the record is listed, not on the actual industry of the business. As a result, an outdoor information centre which happens to sell books and is listed in the "Book - Retail" section of the Yellow Pages would have the same ASIC code as a regular book store. Another problem with this source of data is the fact that many Yellow Pages business classifications correspond to more than one ASIC code. This implies that there can be more than one land-use designation assigned to an individual business. Unfortunately, there is no automatic way of deciding which ASIC code is more appropriate for the particular business concerned. This is further complicated by the fact that some businesses have more than one entry in the Yellow Pages and potentially even more associated ASIC codes.

Another weakness of the TBF is that it provides information only on businesses that are listed in the Yellow Pages. Although instinctively one might assume that most businesses of substantial size (which have their own phone number) are bound to be found in it, experience has shown that some relatively large and popular commercial institutions are actually not listed. For this reason the TBF cannot be considered as a comprehensive source of land-use locations.

Furthermore, the information contained in the TBF was derived from the printed version of the Yellow Pages. Any existing error in the Yellow Pages would therefore be carried over, and additional errors may have been added during the transferral process. Some of those errors follow a pattern and were identifiable and correctable, but others appeared to be random. Errors of the latter type were very hard to detect, but so far, through spot checks, less than 1% of the data was identified as containing such errors. A crucial type of error was one that occurred in the address information since the process of transforming the tabular data into a spatial database required exact and correct full addresses. Unfortunately, it was found that a substantial number of records do not have exact and correct full addresses, rendering them unusable.

The Journey to Work data

The Journey to Work (JTW) data contains information on employed persons from the 1991 Census of Population and Housing (ABS 1991). It includes substantial information about the location (in terms of destination zones which are subsets of SLAs) and the industry (in terms of ASIC Division) of a person's main place of employment.

Considering that it comprises all employed persons on the census day (salary and wage earners or otherwise) and can be classified by geographic location (SLA) and land-use designation (ASIC Division), the data can be used to generate a more comprehensive database of employment (actually based on employed persons) by SLA and ASIC Division than the Business Register. It also includes employment numbers for agricultural units that are missing from the Business Register. Its main drawback is that it is only updated as often as the census is undertaken, that is every five years. Moreover, as with the Business Register, its smallest 'useable' geographic unit is the SLA. Although it is also subdivided into destination zones, these geographic units are not commonly used in other statistics, making them practically unusable.

COMBINING THE THREE DATA SETS

Because of the shortcomings of the individual data sources, it was decided that information from all of them should be combined to satisfy the requirements of the land-use activity database. First, the employment information from the JTW data would be used to correct Business Register employment numbers in each SLA, and these resulting numbers would then be the basis for the measure of activity levels. Then, to disaggregate the data into smaller geographic units, the TBF data would be used. Disaggregation to the level of individual businesses was, however, not possible because of the significant number of missing businesses in the TBF. Instead, the Census Collector District (CCD) was chosen. The CCD is an immediate subdivision of the SLA and the

smallest geographic unit used by the Australian Bureau of Statistics. It ranges in size from 500 sq metres in the central city area to 290 sq km in the outskirts. For each SLA, the employment number in each ASIC Division would be distributed among its constituent CCDs based on the proportion of existing individual businesses in each ASIC Division as calculated using the TBF data.

A common land-use designation was needed before combining the three data sets. Although both the Business Register and the TBF use ASIC designations to the lowest 4-digit 'Class', it was considered impractical and risky to work at such high level of detail. It was impractical because the JTW data was not available at that level and because there are over 500 ASIC Classes, and it was risky because of the problems associated with questionable ASIC designation in the TBF data set, as explained earlier. Instead, it was decided that a more manageable set of land-use designations was to be used, and the ASIC 'Division' level, consisting of 13 industry designations, was selected (see Table 1).

Table 1 ASIC Division Classifications

ASIC Division Code and Description	
A	Agriculture, Forestry, Fishing & Hunting
B	Mining
C	Manufacturing
D	Electricity, Gas & Water
E	Construction
F	Wholesale & Retail Trade
G	Transport & Storage
H	Communication
I	Finance, Property & Business Services
J	Public Administration & Defence
K	Community Services
L	Recreation, Personal and Other Services
M	Non-Classifiable Economic Units

Furthermore, by using ASIC Divisions instead of ASIC Classes, the problem of unavailable employment numbers in SLAs with five or less business locations may be eliminated because in aggregation (of several classes into a division) it would be more likely that all SLAs contain more than five business locations.

DATA PREPARATION

The Business Register

The Business Register data which was initially in tabular form was transformed into a spatially oriented database by merging its entries with associated SLA boundaries from a MapInfo™ GIS map. The tabular data was matched with the map using information common to both, namely SLA name. The resulting product was a database that contains both the information from the Business Register and the geographic SLA boundaries from the map, all stored within a single set of interrelated GIS files. From this GIS database, any subset of the data can be extracted based on any of the available attributes (such as land-use designation or area coverage) and can be graphically presented on a map. As an example, the GIS package MapInfo™ was used to generate Figure 1 which shows the employment in each SLA in the Melbourne central area for ASIC Division F (Wholesale & Retail Trade).

In this map, every 10 employment units in an SLA was represented by a dot that is randomly located within the SLA. The resulting concentration of dots in an SLA suggests the employment density in that SLA.

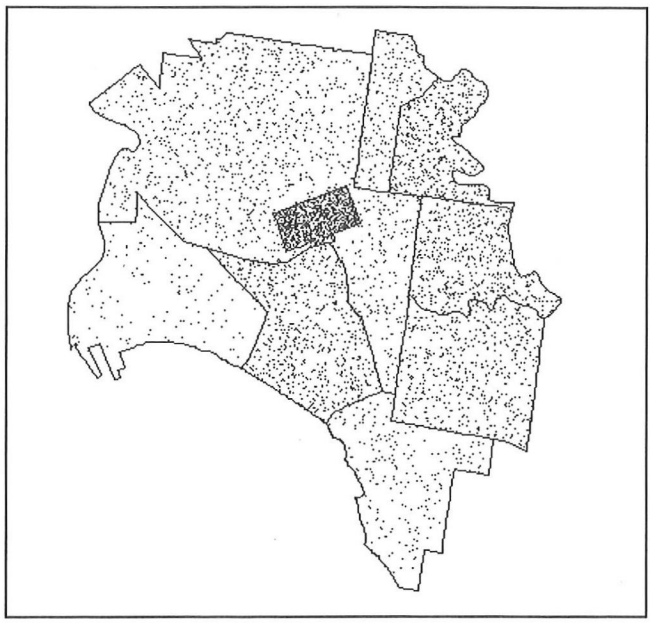

Figure 1 SLA employment density in the wholesale and retail trade industry in Central Melbourne

The TBF

After extracting the data from the TBF CD-ROM and converting it into a tabular format, the data was transformed into a GIS database through a process called geocoding. The process, which is inherent to MapInfo™, translated street addresses into longitude-latitude coordinates that can be read and represented as a point on a map. This was done by matching the street addresses from the TBF with street address ranges in a comprehensive database of street segments. If a match were found, longitude-latitude coordinates were then calculated for the TBF record based on its approximate position relative to the two extremes of the street segments whose coordinates are known. The process requires correct street addresses to succeed. At the first run only 69% of all the records, the ones that had correct complete addresses (in terms of either street number and street name or cross streets), were successfully geocoded. Through several stages of data editing, an additional 5% were geocoded. With the remaining 26%, a systematic search was conducted of addresses that contained some types of landmark information, such as a shopping centre or building name. This information was used to geocode a further 2%, bringing the total geocoded entries to 77% (figures do not add up due to rounding). Unfortunately, the geocoding success rate was not uniform throughout the metropolitan area. It ranged from over 90% in the city centre to less than 1% in some outer suburbs. This was primarily due to the level of information available on the street segments used for geocoding. Less information was available for the outer suburbs. However, if only the suburbs with a geocoding rate of 60% or more were considered, as much as 70% of all entries (only 7% less than the total) would be included.

The subsequent problem to overcome was that of multiple ASIC designations. Some entries, as mentioned earlier, had more than one land-use designation. This was either because their Yellow Pages classifications corresponded to more than one ASIC code, because they had more than one entry in the Yellow Pages, or because of a combination of the two. For those entries, it was decided that they should be considered as single businesses having land-use designations that are proportionally distributed according to the number of occurrences of ASIC codes assigned to it. For example, an entry that has been designated as belonging to ASIC Division C on one count and ASIC Division F on three counts would be computed as being 1/4 of a business in Division C and 3/4 of a business in Division F.

ASIC Division F on three counts would be computed as being 1/4 of a business in Division C and 3/4 of a business in Division F.

COMPARISON OF BUSINESS REGISTER AND TBF

To justify the use of the TBF to disaggregate the Business Register data, it was necessary to first verify that the TBF data was indeed accurately distributed geographically. A comparison was therefore made to check whether the TBF and Business Register gave similar spatial distributions of employment sites at the SLA level. If so, it could be assumed that the distribution was accurate even at the lower CCD level. As mentioned earlier, it was only possible to geocode (and use) 77% of the TBF data, but even initially (before geocoding) many businesses were found to be missing from the TBF. In many cases, this may be a random phenomenon, but it may also be due to a low entry rate in the Yellow Pages by certain types of businesses or by businesses in particular locations. In the latter cases, the non-appearance of those businesses may not only affect their population size but also their proportional geographic distribution. It may even be that the distribution is erroneous because the total population itself is too low. Whatever the reason, the use of the TBF data as a basis for distributing Business Register employment data would then be less justifiable. Table 2 shows a comparison of the total number of business locations available in the Business Register and TBF data for each land-use designation (ASIC Division), while Figure 2 shows the relationship between the total number of business locations in each SLA in both data sets.

Table 2 Number of business sites for each ASIC Division in the Business Register and TBF data

	ASIC Division	Number of Sites		Ratio
		Business Register	TBF	
A	Agriculture, Forestry, Fishing & Hunting	3,691	271	0.07
B	Mining	272	106	0.39
C	Manufacturing	13,167	14,188	1.08
D	Electricity, Gas & Water	224	56	0.25
E	Construction	15,128	7,736	0.51
F	Wholesale & Retail Trade	37,867	33,236	0.88
G	Transport & Storage	5,923	2,485	0.42
H	Communication	824	65	0.08
I	Finance, Property & Business Services	33,891	17,048	0.50
J	Public Administration & Defence	872	1,331	1.53
K	Community Services	18,255	12,531	0.69
L	Recreation, Personal and Other Services	11,611	10,554	0.91
	TOTAL NUMBER OF SITES	141,774	99,607	0.70

From Table 2, it can be seen that in most cases the number of Business Locations in the TBF is lower than in the Business Register. As explained earlier, this is partly due to the fact that many businesses do not have a listing in the Yellow Pages and partly because of the less than perfect rate of geocoding. Given that only 77% of the TBF sites were geocoded, it can be seen that the TBF sites are still under-represented since the ratio of total TBF sites to Business Register sites is only 70%. In the case of ASIC Division C (Manufacturing), the slightly higher number in the TBF suggests the problem of non-uniform definitions of industry by the two data sources. A hot bread shop, for example, is consistently defined as a manufacturing business by TBF. Under the ASIC system, however, there exists a different industry classification for bread manufacturers (Manufacturing) and bread sellers (Retail Trade). As the Business Register assigns only one ASIC industry classification to a site according to its main line of business, it may well be that in the Business Register a bakery is classified as retail business instead. As for the large discrepancy in the J (Public Administration & Defence) land-use designation, this is due to the many divisions within some public offices. For example, departments in a municipal office may be individually listed in the TBF. These different departments are part of a single 'business location' as defined by

the Business Register, but in the TBF, because of their individual names, they are considered separate entities.

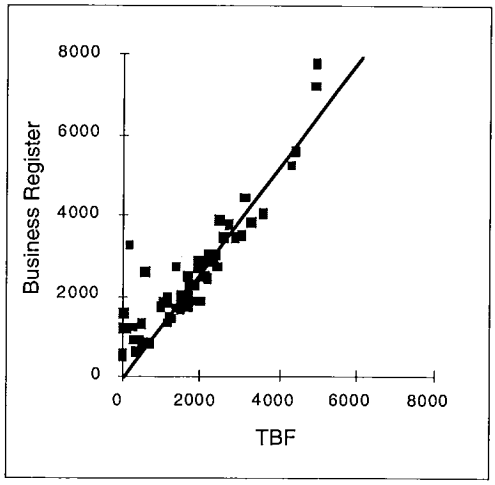

Figure 2 Business location spatial distribution (by SLA) from Business Register and TBF (each data point represents one SLA)

It can be seen from Figure 2 (as shown by the line) that there is generally a positive relationship between the number of employment sites in an SLA in the Business Register and the corresponding data in the TBF (after allowing for the fact that only 77% of TBF sites were geocoded). At the upper end, the Business Register shows more employment sites in the top two SLAs (the Melbourne central city areas), while at the lower end there is also an over-representation in the Business Register, mainly because of the discrepancies already noted in Table 1 for ASIC Divisions A, B, D and H.

As the Business Register contains a more complete set of business locations by SLA than the TBF, the spatial distribution of TBF data at the SLA level can be validated by comparing it to that of the Business Register. Using a GIS-aided procedure, the TBF data was first aggregated to the SLA level. Then, for each of the land-use types (ASIC Divisions), the proportional distributions of the number of businesses in all SLAs were calculated for both the TBF and the Business Register. These proportional distributions were then compared using the correlation coefficient formula:

$$\rho_{TBF,BR} = \frac{\mathrm{cov}(TBF,\ BR)}{\sigma_{TBF} \bullet \sigma_{BR}}$$

where

$$\sigma^2_{TBF} = \frac{1}{n}\sum_{i=1}^{n}(TBF_i - \mu_{TBF})^2 \quad \text{and} \quad \sigma^2_{BR} = \frac{1}{n}\sum_{i=1}^{n}(BR_i - \mu_{BR})^2$$

These correlation coefficients were calculated with and without windsorising the data. In other words, two sets of correlation coefficients were calculated, first using all the data, and then only for data that fell within 3 standard deviations of the mean (in either the TBF or the Business Register distribution). Figures A-1 to A-12 in the Appendix show the relationships between the two distributions for each land-use designation. The thirteenth land-use designation, M (Non-Classifiable Economic Units), was not considered since the TBF did not contain any data in that category. In each of those figures, the circled points were the ones outside the range of three standard deviations from the mean and therefore omitted from the calculation of the second set of

particular case, the correlation coefficient was much higher before windsorising because of the effects of one single point that was substantially far from the mean value (see Figure A-2). In one other case, land-use designation H (Communication), the correlation coefficient actually increased after windsorising but not substantially.

Table 3 Correlation coefficients between TBF and Business Register proportional distribution of business locations (in SLA)

	Land-Use Designation	Correlation Coefficients	
		Before Windsorising	After Windsorising
A	Agriculture, Forestry, Fishing & Hunting	-0.10	-0.29
B	Mining	0.90	0.41
C	Manufacturing	0.92	0.86
D	Electricity, Gas & Water	0.09	0.09
E	Construction	0.80	0.75
F	Wholesale & Retail Trade	0.93	0.93
G	Transport & Storage	0.87	0.77
H	Communication	0.54	0.62
I	Finance, Property & Business Services	0.95	0.94
J	Public Administration & Defence	0.90	0.56
K	Community Services	0.94	0.92
L	Recreation, Personal and Other Services	0.93	0.87

A perfect correlation is a situation where in both sets of data the values are spatially distributed in the same way. The coefficient of correlation in such case would have the value of 1. Where there is little relation between the two sets of data, that is, where the value in one set does not have any significance on the corresponding value in the other set, the correlation coefficient would be close to 0. A negative correlation suggests that the value in one set has a negative association on the value in the other set. In other words, a high value in one set corresponds to a low value in the other set, and vice versa.

A correlation coefficient of 0.85 or higher was considered sufficiently high to exhibit a matching distribution. On this basis, it can be seen from Table 3 that five land-use designations, namely C (Manufacturing), F (Wholesale & Retail Trade), I (Finance, Property & Business Services), K (Community Services) and L (Recreation, Personal and Other Services) would be regarded as having matching spatial distributions in the TBF and Business Register data. It can be calculated, from Table 2, that these five land-use designations comprise 81% of all business sites according to the Business Register or 88% according to the TBF.

As for the rest, one of the reasons for their low correlation coefficients was that the TBF and Business Register were describing different populations, as illustrated by the ratios of the number of sites in each ASIC Division in each data source (Table 2).

COMPARISON OF BUSINESS REGISTER AND JTW

The previous section has shown that for five major ASIC Divisions the distribution of employment sites in the Business Register reasonably matches that of the TBF. In this section a comparison is made between the number of employed persons in the Business Register to that of the Journey-to-Work (JTW) data.

The two data sets do not contain exactly the same information. The Business Register records employment positions, including part-time positions, of wage and salary earners, while the JTW data contains information on main employment of employed people only. The comparison of the total employment in each ASIC Division in the Business register and the JTW data is shown in

data contains information on main employment of employed people only. The comparison of the total employment in each ASIC Division in the Business register and the JTW data is shown in Table 4. In addition to what is shown in the table, the JTW data also includes about 12,000 jobs for which no specific employment category was given.

Table 4 Employment in each ASIC Division in the Business Register and JTW data

	ASIC Division	Employment		Ratio
		Business Register	JTW	
A	Agriculture, Forestry, Fishing & Hunting	6,510	8,570	1.32
B	Mining	1,477	2,298	1.56
C	Manufacturing	269,675	225,962	0.84
D	Electricity, Gas & Water	14,804	13,177	0.89
E	Construction	70,251	63,181	0.90
F	Wholesale & Retail Trade	273,915	250,033	0.91
G	Transport & Storage	60,672	55,829	0.92
H	Communication	29,658	26,168	0.88
I	Finance, Property & Business Services	228,721	159,358	0.70
J	Public Administration & Defence	58,328	62,978	1.08
K	Community Services	236,012	221,976	0.94
L	Recreation, Personal and Other Services	100,438	73,513	0.73
	TOTAL EMPLOYMENT	1,357,513	1,175,607	0.87

Figure 3 shows the relationship between the total employment numbers, spatially distributed by SLA, for the two data sets. The two distributions have a correlation coefficient of 0.99 (0.97 after windsorising the two circled points which are again values that correspond to the central city area) which indicates that the Business Register and the JTW data have very similar spatial distributions of employment. The line in Figure 3 indicates the average ratio of 0.87 for JTW employment compared to Business Register employment numbers.

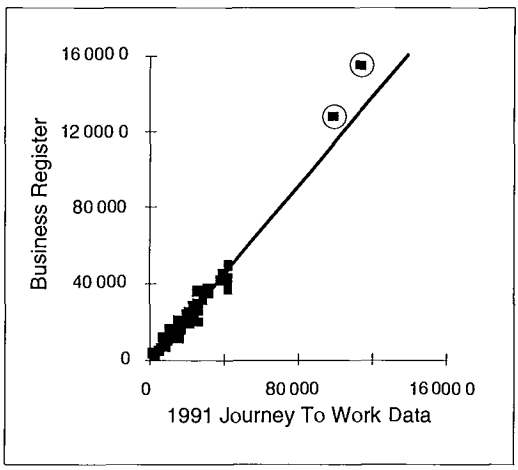

Figure 3 Plot of Journey to Work against Business Register employment spatial distribution (by SLA)

DISTRIBUTION OF EMPLOYMENT BY CCD

The previous sections have shown that the comparisons between the distribution of employment sites in the TBF and the Business Register and between the distribution of employment in the Business Register and the JTW data gave reasonable results. Given that, this section will explore the process of distributing the Business Register number of employment by ASIC Division in each SLA to the CCDs within the SLA using the spatial distribution of employment sites obtained from the TBF.

Prior to the analysis of the data, it was thought that the JTW data would have a more complete figure for employment, especially since it includes people who are not wage or salary earners. However, as seen in Table 4, the total JTW employment was only 87% of the Business Register employment numbers, even after considering the 12,000 uncategorised jobs. Whether this was because the JTW data was restricted to one job per person, or because the Business Register contained many businesses which no longer exist, is not clear. This will be investigated in more depth in the future. At any rate, it was felt inappropriate to scale the Business Register data to match the JTW data at this stage.

For the five ASIC Divisions whose TBF data spatial distributions matched that of the Business Register, the proportions of business locations in each CCD within each SLA were calculated. These proportions were then used to distribute the employment numbers in each SLA (from the Business Register) to the respective CCDs. It is these numbers that are used as a measure of land-use activity. As an illustration, Figure 4 shows the same information as in Figure 1 but disaggregated to the CCD level. As with Figure 1, each dot represents 10 employment units randomly spread in the geographic unit area (the CCD in this case). It can be seen that Figure 4 provides a much better description of the spatial distribution of retail and wholesale sites than Figure 1. It can be seen that within the CBD area (the rectangular area in the middle of the figure) the distribution of retail and wholesale employment is not uniform but concentrated in the central part, as is in fact the case (this is the area between Swanston Street and Queen Street). Also, other CCDs which are known to be parklands show little, if any, retail and wholesale employment as expected.

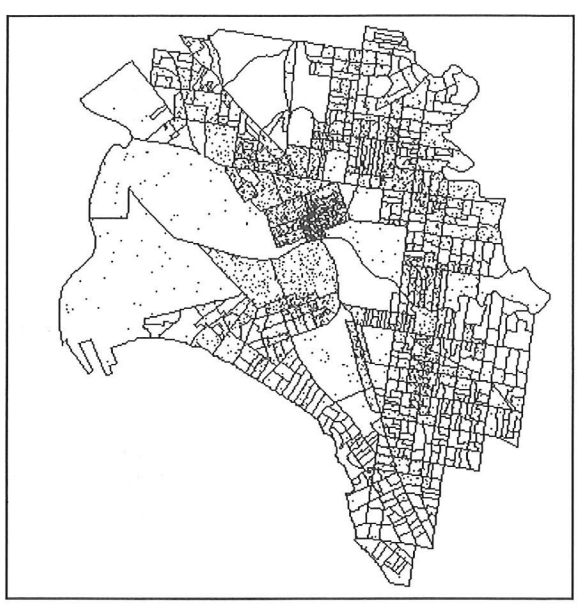

Figure 4 CCD employment density in the wholesale and retail trade industry in Central Melbourne

RESULTS

A disaggregate land-use activity database was produced for all land-use designations and for the whole Melbourne metropolitan area. However, because only five land-use designations were determined to be properly spatially distributed, only for those land-uses can the database be considered reliable. Still, they comprise over 80% of all business locations and employment. Also, because of the lower rate of geocoding success in some outer suburbs, the results for those areas are questionable. Nevertheless, 70% of all business locations (only 7% less than the total geocoding success rate) are within a suburb whose success rate was over 60%. For those latter suburbs the resulting database can be used confidently.

In addition, some useful databases other than the disaggregate land-use activity database were also produced. One of the intermediate results, for example, is a GIS form of the Business Register data. This is in fact a considerably useful land-use activity database in itself. Its only shortcoming is its relatively large geographic unit (ie the SLA). Still, for many different applications it can be useful; this is particularly so because it inherently contains a measure of activity (employment size) and because it can be spatially manipulated (GIS-based).

The geocoded TBF data is, by itself, also a very useful database. Although it is far from comprehensive and lacks any measure of activity, it can be useful in other ways. For one thing, as it contains the names of individual entries of business units, it can serve as a database of landmarks whose exact locations are known and are electronically recorded. For other components of the VITAL toolkit, this information is very valuable. It can, for example, be used for determining the exact geographic point of the origin or destination of a trip, if that point can be related to be a business that has an entry in the Yellow Pages (and can be geocoded).

CONCLUSIONS

Although a totally disaggregate (to the level of individual location) land-use activity database was not achieved, the results for some of the more significant land-use designations and for most part of the Melbourne metropolitan area have been quite satisfactory. The sizes of CCDs, particularly in the central area of the city, are small enough for the data to be considered highly disaggregate. For the land-use designations that remained at a high level of spatial aggregation and for the suburbs where the geocoding rate of the TBF was low, it would be desirable to find other means of disaggregation. However, even at the level of aggregation provided by the Business Register, the data can be highly effective. What still needs to be done is to resolve the anomalies found in the data and to obtain more justifiable factors for correcting the employment levels in the Business Register.

REFERENCES

ABS (1991) *1991 Census Dictionary*, Australian Bureau of Statistics, Canberra.

ABS (1992) *Explanatory Notes for Business Register Data*, Australian Bureau of Statistics, Canberra.

ITE (1987) *Trip Generation, 4th Edition*, Institute of Transportation Engineers, Washington DC.

Richardson, A.J. and E.S. Ampt (1993) The Victorian Integrated Travel, Activities and Land-Use Toolkit, *Vital Working Paper*, VWP93/1, Transport Research Centre, Melbourne.

Telecom Australia (1992) *Telecom Business Finder User Guide*, Australian and Overseas Telecommunications Corporation, Melbourne.

APPENDIX

Proportional distribution of business locations by SLA in the business register and TBF database

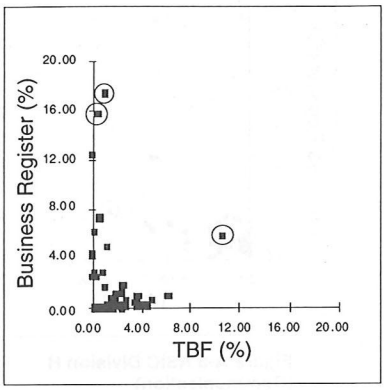

Figure A-1 ASIC Division A (Agriculture, Forestry, Fishing & Hunting)

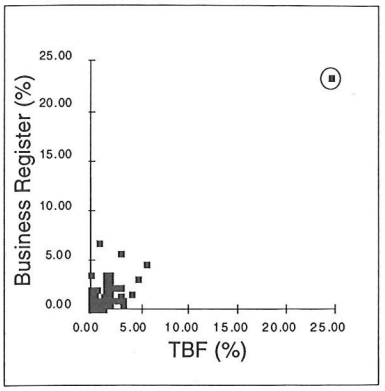

Figure A-2 ASIC Division B (Mining)

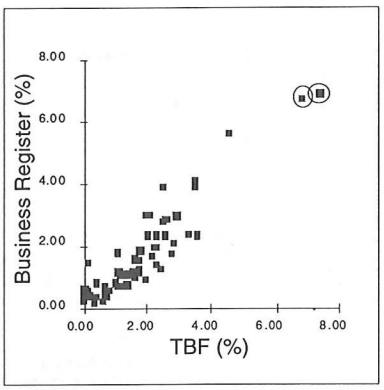

Figure A-3 ASIC Division C (Manufacturing)

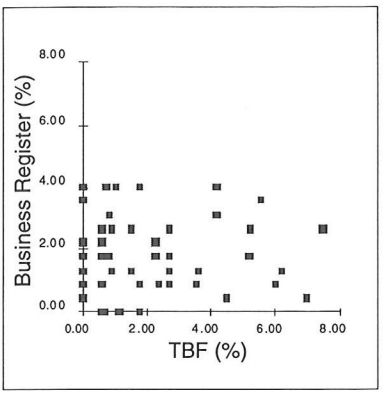

Figure A-4 ASIC Division D (Electricity, Gas & Water)

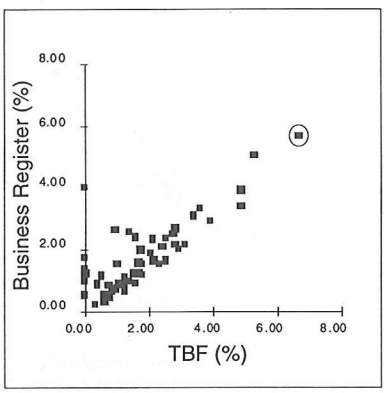

Figure A-5 ASIC Division E (Construction)

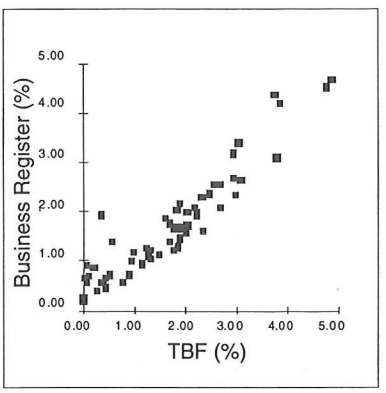

Figure A-6 ASIC Division F (Wholesale & Retail Trade)

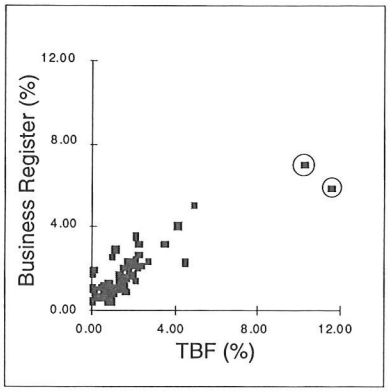

Figure A-7 ASIC Division G (Transport & Storage)

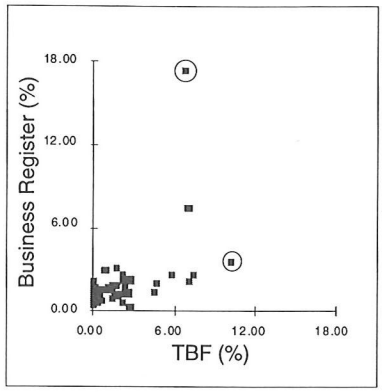

Figure A-8 ASIC Division H (Communication)

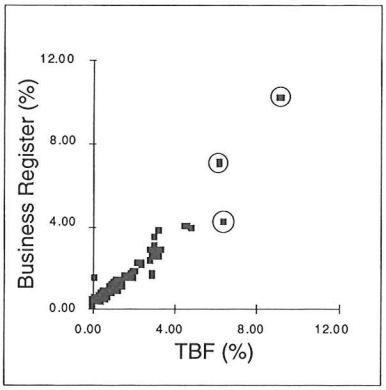

Figure A-9 ASIC Division I (Finance, Property & Business Services)

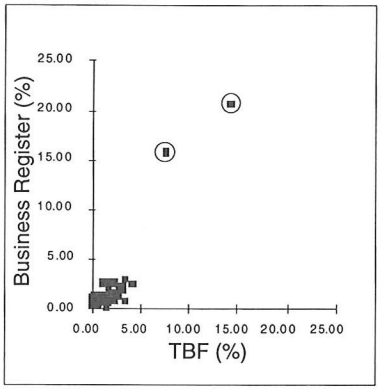

Figure A-10 ASIC Division J (Public Administration & Defence)

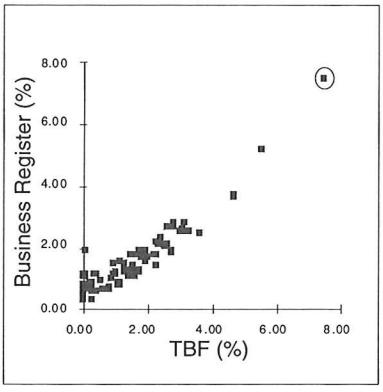

Figure A-11 ASIC Division K (Community Services)

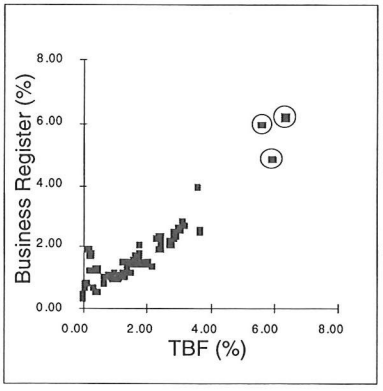

Figure A-12 ASIC Division L (Recreation, Personal & Other Services)