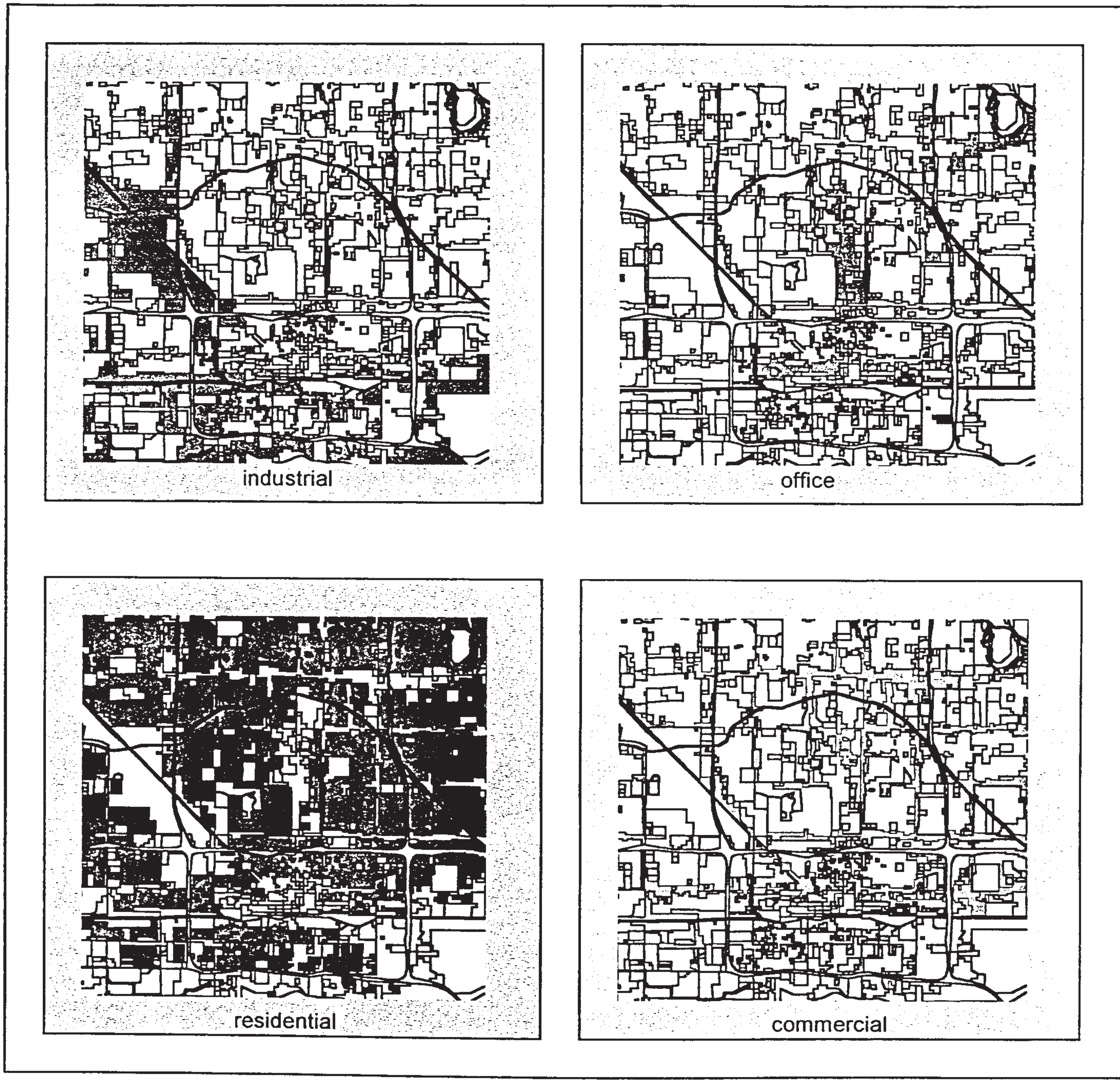
GIS ANALYSIS AND DATA ENHANCEMENT STUDY

Prepared for
The Maricopa Association of Governments



Submitted by
Barton-Aschman Associates, Inc.
in Association with AMPG, Inc.

January 1996

Contents

Contents (Continued)

Chapter (Continued)

Tables

Contents (Continued)

Tables (Continued)

Figures

Contents (Continued)

Figures (Continued)

Executive Summary

In cooperation with local governments in the Phoenix metropolitan area, the Maricopa Association of Governments has the responsibility for regional land use, transportation, and air quality planning. MAG operates a number of programs to fulfill those reponsibilities. MAG, for example, is responsible for land use forecasts for the region. Also, the agency has developed and maintains travel forecasting capabilities which are used to support transportation programs in the region as well as to support air quality conformity analysis required by the Intermodal Surface Transportation Efficiency Act (ISTEA).

MAG's ongoing planning program frequently includes elements to improve its modelling capabilities, to make them more effective to better serve the region. As part of this ongoing initiative, the GIS Analysis and Data Enhancement Study was started in August, 1994. The three principal objectives of this project were:

- **Transportation Network Database Improvements:** A major effort to overhaul MAG's travel forecasting models was undertaken in 1993 and 1994, part of which involved the transfer of the models from the antiquated UTPS[1] software environment to EMME/2[2] operating on unix workstations. The models were also recalibrated. The GIS Analysis and Data Enhancement Project addressed additional needs to improve the transportation network databases which support the travel forecasting model.

---

[1] UTPS, or Urban Transportation Planning System, is an old modeling system originally developed by USDOT and operates in a mainframe computing environment. It is no longer supported by the federal government.

[2] EMME/2 is a proprietary travel forecasting package developed and distributed by INRO, Inc. It operates on personal computers and unix workstations.

- **DRAM/EMPAL Feedback Loop Mechanism:** MAG's land use forecasting process relies heavily on DRAM/EMPAL, a land use forecasting model developed by Steven Putnam at the University of Pennsylvania. The GIS Analysis and Data Enhancement Project introduced a substantial improvement in the way this model is operated. This improvement addressed the need to assure that land use forecasts remain consistent with the measures of accessibility on which they are based.

- **Subarea Allocation Model:** DRAM/EMPAL generates land use forecasts for a large scale geography -- regional analysis zones. MAG planners must disaggregate these forecasts to smaller scale geographies, such as the traffic analysis zone system. Also, MAG responds to frequent demands from member governments for forecast information on other geographies. Another objective of the GIS Analysis and Data Enhancement project was to develop a GIS-based method for disaggregating forecast information.

Transportation Network Database Improvements

Transportation networks play an important role in the simulation of travel demand. MAG maintains transportation network databases for the years 1993, 1995, 2005, and 2015. These networks reflect the gradual development of additional transportation infrastructure consistent with the regional transportation plan (freeway extensions, roadway widenings, etc.). Originally, these networks were coded according to an arbitrary coordinate system. While this was adequate to allow travel forecasting packages (such as EMME/2) to display networks onscreen, the network maps could not be combined with other sources of geographic information, which is all coded according to Arizona Central State Plane (ACSP) coordinates. The principal objective of the network improvments implemented in this project was the conversion to state plane coordinates.

ARC/INFO was used to convert MAG networks from their original coordinate system to ACSP coordinates. The street centerline coverage for Maricopa County was used as the basemap for the conversion process. MAG had previously devoted considerable time to the development of the street centerline map, which is considered to be the most accurate representation of street features in Maricopa County.

The conversion involved a series of steps to improve the positional accuracy associated with the travel networks, as follows:

- EMME/2 networks were imported into ARC/INFO. Formulas were derived to translate and scale the networks to conform generally with ACSP coordinates. Once accomplished, the networks could then appear onscreen in the same coordinate space as the underlying street network they represented.

- The networks were then adjusted to match the street centerline basemap through a GIS process known as *rubber-sheeting*. Rubber sheeting essentially *warps* maps (such as the network map) to match a background basemap (such as the street centerline map).

- The rubber-sheeting process did not prove to be an effective technique for freeways, frontage roads, and interchanges because of the lack of shape points in EMME/2 networks. Consequently, they were adjusted manually. An application to accomplish this was developed. In addition, another application to automatically migrate geographic revisions to freeways in one network to other networks was developed to reduce the time involved with editing.

Other revisions to MAG networks were made as well. Nodes in the networks were renumbered according to a new numbering scheme established by MAG. Based on conflation, node numbers in the base network were migrated to other future year networks to establish a consistent representation. Transit lines descriptions and turn penalty tables were modified to reflect the new numbering scheme.

Finally, continuing work that was started in a previous study, revisions to the traffic analysis zone (TAZ) geography were made. Two versions were created, both providing additional geographic detail as well as expanding the modeling area to cover all of Maricopa County. One TAZ geography resulted in 1,558 TAZs, thereby complying with MAG's current 1,600 zone limit imposed by the EMME/2 license. The other TAZ geography resulted in 1,760 TAZs and would be applicable in the future when MAG increases its EMME/2 license provision.

DRAM/EMPAL Feedback Loop Mechanism

The second major objective of the GIS Analysis and Data Enhancement Project addressed the need for revisions to the DRAM/EMPAL forecasting process.

DRAM/EMPAL generates forecasts of land use for regional analysis zones (RAZs) based on a number of input variables, including countywide control totals, estimates of the availability of developable land, and other factors. *Accessibility* is another important factor which drives DRAM/EMPAL. Measures of accessibility typically come from the travel forecasting model (for the previous forecast year). Critics have questioned whether the resulting land use forecasts remain consistent with the level of accessibility on which they are based.

Putnam, the developer of DRAM/EMPAL, and Florian, the developer of EMME/2, recommend the implementation of a *feedback loop mechanism*. The feedback loop mechanism calls for an iterative process whereby transportation system accessibility is reestimated based on the land use forecast generated by DRAM/EMPAL until a unique closed solution is determined. This method has the effect of attenuating forecasts to a solution which is consistent with the measures of accessibility on which it is based.

As a practical matter, the implementation of the feedback loop mechanism requires a program to cyclically drive both DRAM/EMPAL and the EMME/2 model chain. Land use forecasts generated by DRAM/EMPAL are converted to a form required by EMME/2, which is then used to generate a new estimate of travel time. This new estimate of travel time is then converted for use by DRAM/EMPAL, which is called on to generate a new land use forecast. This process continues until closure is achieved.

The DRAM/EMPAL feedback loop mechanism was implemented as ARC/INFO AML program called ***tlume***. A user-friendly interface was created to provide input menus and forms to control the process. On the DRAM/EMPAL side, the ***tlume*** application essentially replaced a series of ***DOS*** and ***Paradox*** scripts that formerly drove the DRAM/EMPAL forecasting process when it resided on personal computers. The ***tlume*** application is more user-friendly than its ***DOS*** predecessor because it makes heavy use of graphic user interface (GUI) capabilities of ARC/INFO that are not available to ***DOS***.

Key elements of the process were designed to be as follows.

- A trip generation dataset for EMME/2 must be derived from intermediate DRAM/EMPAL forecast runs. This raises two issues: (1) the need to disaggregate forecasts of households and employment and (2) the need to generate estimates of a number of other variables employed in trip generation that do not come from DRAM/EMPAL.

- An abbreviated form of the MAG travel forecasting process running on EMME/2 is needed. Putnam argues that PM peak hour highway accessibility drives land use decisions, consequently the travel forecasting process need only focus on this aspect of travel demand.

- Travel time estimates generated by EMME/2 on a TAZ basis need to be aggregated to RAZs.

DRAM/EMPAL Data Disaggregation

The program ***tazdata.c*** creates a trip generation dataset from intermediate DRAM/EMPAL results. Key elements of the procedure are:

- Household and employment estimates from DRAM/EMPAL are disaggregated from RAZs to TAZs. This is accomplished by factoring TAZ-level household and employment totals for a previous forecast year to match the new RAZ-level control totals.

- Estimates for other variables, such as population, are derived from information coming from DRAM/EMPAL. Estimates for special population groups (transient, seasonal, and group quarters population) are derived from countywide control totals for the forecast year in question.

- Network related information, such as terminal times, parking costs, distance to the ASU campus, etc., remain the same as they were for the previous forecast year.

Abbreviated Travel Forecasting Process

For the most part, the travel forecasting process embedded in ***tlume*** calls on standard EMME/2 macros in use at MAG, so any changes to the regular forecasting process implemented at MAG would be immediately reflected in DRAM/EMPAL. However, the process does involve several special techniques not present in the regular MAG model chain:

- Mode choice is not run. Instead, the mode choice module is replaced in ***tlume*** by a vehicle occupancy model which directly converts person trip estimates to vehicle trip estimates for assignment. Typical prevailing mode choice characteristics apparent in MAG regional model runs are reflected in the vehicle occupancy model.

- The ***tlume*** implementation concentrates only on PM peak hour assignments. Other time periods are of no interest to DRAM/EMPAL.

These revisions make the EMME/2 process operate much faster on a much smaller databank than is normally associated with MAG travel forecasts.

Travel Time Estimates

MAG's original ***ttimes*** program used to generate travel time estimates for DRAM/EMPAL was completely rewritten. The new version, ***ttimes2.c***, employs a much improved algorithm with the following features:

- Travel times between RAZs are estimated as the average travel time between TAZs that are members of the RAZ pair in question. The average travel time is weighted by the trip volumes involved in the interchange.

- Travel times between RAZs which are external to the modeling area are estimated in parts -- the part involving travel internal to modeling area is estimated as above -- the part involving travel outside of the modeling area is estimated on airline distance.

Inspection of the results generated by ***ttimes2*** look very plausible and more reasonable than the results generated by the original model, especially for RAZs external to the modeling area.

Closure

The last issue in the feedback loop implementation concerns *closure*. Putnam and Florian argue that a unique and consistent solution can only be achieved when intermediate traffic *volume* results (not travel *times*) are weighted by a convex function. This means essentially that intermediate results for each iteration through the feedback loop should be increasingly discounted with each new iteration. This method was implemented in ***tlume***.

Subarea Allocation Model (SAM)

The last major objective of the GIS Analysis and Data Enhancement project concerned the need for a subarea allocation model to disaggregate DRAM/EMPAL results available from the model for RAZs to more detailed geographies (for TAZs, for example, so they can be used in travel forecasting). Other objectives of the SAM were:

- MAG responds to requests for socio-economic forecast information for a wide variety of geographies, such as water district or municipal boundaries. Conflicts between these geographies can make the development of such data sets extremely difficult.

- Communities in Maricopa County are interested in implementing focused travel models. Focused travel models are derived from the regional model, but provide considerable detail within local communities. MAG will have to respond to demands for disaggregated datasets to support trip generation in focused model areas.

The approach to the data managment problem created by these demands lead to the concept of the *Minimal Analysis Zone* (MAZ) system. In the future, MAG data will be maintained in a highly disaggregate form such that demands for summaries on other geographies can be met through a simple aggregation procedure. The implementation of the MAZ makes heavy use of ARC/INFO's GRID capabilities. GRID is a *raster* representation of geography rather than the more conventional *vector* representation of geography -- the raster representation of geography is analagous to bitmap imagery. This representation of geography facilitates algorithms that are not possible with conventional vector representation and can be inherently more powerful.

The SAM model, implemented in an ARC/INFO AML application called ***sam***, is designed to disaggregate forecast information from DRAM/EMPAL into a highly detailed and disaggregate form. It does this essentially by *simulating* or *emulating* land use decisions made by developers. The key elements are as follows:

- A base year land use coverage has been created which is populated with information about households and employment. 1993 was taken to be the base year. In addition, the base year land use coverage contains information about special population groups (motel rooms, mobile home park locations, nursing homes, etc.) which play a major role in the trip generation.

- The SAM model allocates *growth*. Growth estimates for households and employment come from DRAM/EMPAL for RAZs in comparison with whatever dataset is being used as a base year. Growth estimates for other land uses (e.g., special population groups) come from countywide control totals.

- Lands which are *eligible* for development are identified by the model. The criteria by which land is eligible for development include (1) land is represented as *vacant* and *developable* in the existing land use cover and (2) land is appropriately designated for development in the general plan cover.
- Lands which are eligible for development are then evaluated for growth potential through a scoring system which reflects locational preference criteria. A wide variety of criteria can play a role in judging the growth potential for land, including (1) proximity to the urban area, (2) proximity to major arterial highways, and (3) proximity to other developed land. Special characteristics, for example the propensity for motels to be built almost exclusively around freeway interchanges, can be reflected in the scoring system.
- Growth is then allocated to the highest ranking lands until it is completely absorbed. Growth is absorbed at density levels associated with the general plan designations, which can be altered as needed to completely absorb growth for any time frame.
- A new forecast land use cover, combining both the existing land use cover as well as lands identified to absorb growth, is then generated.

The SAM model represents a revolutionary approach to disaggregating socio economic information and provides MAG with a framework on which to build more effective land use forecasting techniques in the future. It also provides motivation for improvements to underlying land use and general plan datasets.

Conclusions

The GIS Analysis and Data Enhancement Project successfully achieved a number of important objectives for MAG:

- GIS was used to implement substantial improvements in MAG travel demand forecasting networks. In addition, MAG now has a GIS application to import and export EMME/2 networks into the ARC/INFO environment, allowing forecasting results to be displayed in association with other sources of geographic data.
- The DRAM/EMPAL feedback loop mechanism is now available for use in the next round of land use forecasts to be generated by MAG. This represents the first operational implementation of the feedback loop mechanism in the country.
- The Subarea Allocation Model is now available for use, both to support MAG forecasting efforts in the future as well as to address needs for data disaggregation to support focus model efforts in the region. The SAM represents an advanced and powerful modeling technique of a truly innovative nature.

1
Sub-Regional Allocation Procedures and MAG Data Sources

Over the past several years the Maricopa Association of Governments (MAG) has undertaken efforts to improve and integrate their land use and travel demand forecasting capabilities. As part of this process MAG has acquired and maintains the DRAM/EMPAL modeling suite for land use allocation. In addition, MAG has initiated several major studies to improve the base data upon which these models operate. The EMME/2 travel demand modeling software was recently obtained by MAG to provide a more productive environment for travel demand forecasting. Other efforts have been completed or are currently in process related to improving model components or adding functionality to the travel forecasting process.

Although some of these activities were inspired by the requirements imposed by the Clean Air Act Amendments of 1990 (CAAA90) and the Intermodal Surface Transportation Efficiency Act (ISTEA), MAG has historically focused a great deal of attention on improving processes and methodologies that will allow them to better serve their member agencies, and provide high quality planning information to the policy process for the MAG Region.

MAG has developed a significant Geographic Information System (GIS) capability and has been successful in efforts to integrate a great deal of information and disparate data sources into this environment. The GIS is viewed by many as the most appropriate environment for the analysis and integration of data related to transportation and land use, primarily because most information associated with these processes is inherently locational, and many of the demands for analysis now require answers that can only be provided if spatial relationships are understood. This is particularly true in the area of land use allocation and the analysis of socioeconomic data.

As part of this long term strategy to improve planning capabilities, MAG has contracted with Barton-Aschman Associates, Inc. (BA) to update some of the key variables utilized in the forecasting of socioeconomic data and to develop an improved process for subregional allocation (defined here as allocating relevant socioeconomic variables from the regional level to smaller areas of geography, based on relative opportunity). This is a necessary requirement in the context of a larger objective, which is to integrate the land use and travel demand forecasting processes.

Important aspects of developing a sub-regional allocation capability include; understanding current processes employed by MAG for these purposes, evaluating data and techniques used in other areas, and reviewing MAG data sources in the context of potential subregional allocation requirements.

The purpose of this report is to document efforts made in these areas. The first section contains a general overview of the MAG socioeconomic forecasting process, including a discussion of current practices utilized for sub-regional allocation. Section two provides a review of components and data considered in sub-regional allocation processes employed in other areas. This review is based on a series of contacts with other MPOs and agencies, and focuses on alternative approaches to sub-regional allocation employed by these organizations. This section concludes with a brief discussion of the key components normally considered in sub-regional allocation activities. The last section includes a general review of existing MAG data sources and a profile of GIS coverages and related data. This profile does not represent the entire range of data assets available to MAG. It does represent a summary evaluation of the data sources initially provided to BA for consideration in this task.

1.1 Overview of Existing Processes

It is, perhaps, most useful to discuss issues associated with sub-regional allocation subsequent to a review of the more comprehensive processes currently utilized by MAG for transportation and land use planning. This is necessary to clearly identify the role sub-regional allocation plays in the context of larger attempts to further integrate land use and travel demand modeling practices.

In order to accomplish this objective, this section will focus on the existing practices for travel demand and land use forecasting, the data elements currently required by these processes and recent efforts to improve data and methods used in these processes.

Travel Demand and Regional Allocation

Travel Demand modeling at MAG is performed using the traditional four step process. Trips are generated from zones based on existing and forecasted socioeconomic variables. These trips are distributed to other zones based on relative opportunity utilizing a gravity model. The distributed trips are then assigned a mode of travel and an iterative approach is utilized to assign the trips to the represented street network. One will note that the starting point for this entire process is the forecasting of zone based socioeconomic variables for trip generation.

On the land use side, MAG follows a three tiered process for the forecasting and allocation of socioeconomic variables. Generally speaking, variables are either forecasted to the county level by MAG, or forwarded to the MAG at the county level by other organizations involved in the projections of these characteristics (mostly notably by the Arizona Department of Economic Security). Population projections at the county level are generally produced utilizing cohort-survival and migration models, while employment totals are generated using economic techniques.

These county level projections are allocated to 141 Regional Analysis Zones (RAZ) in Maricopa County utilizing the DRAM/EMPAL land use allocation modeling suite. DRAM and EMPAL are registered trademarks of S.H. Putnam Associates. The DRAM element allocates household locations, while EMPAL allocates employment characteristics. Other elements of the suite calculate land absorption or consumption. Employment is allocated first, based on existing or previous forecast year distributions, land use, and accessibility measures. Households are then distributed based on the distribution of employment. Population is derived from the number of households, but is constrained by the county control totals. One of the critical determinates of the distribution is the measure of impedance from a RAZ to all others. This impedance is provided in terms of zone to zone travel times from the travel demand models.

Allocations of these variables are generally performed at five year intervals starting with the base year and traditionally going out to the 50 year horizon. Based on the RAZ projections for employment, households and population, another allocation is performed to disaggregate the data to the Traffic Analysis Zone (TAZ) level. There are currently 1272 TAZs in the modeling area. It is important to note that this process is driven in part by the need to provide appropriate input variables to the travel demand modeling process. While the process of utilizing DRAM/EMPAL to allocate data from the county level to RAZs can be characterized as a mechanical or automated process (although some users may argue with this qualification), the process of sub-regional allocation in most cases is not.

MAG has a considerable amount of data tied to various levels of geography in the GIS (e.g. existing distributions of variables, vacant land, planned and proposed developments, and general plan land use) which greatly assists this process. However, the current process utilizes manual or pseudo-automated techniques, and local input to derive future estimates of these variables at the TAZ level. In most cases, this involves allocating *change* from existing or prior forecast intervals to individual TAZs based on the RAZ forecasts. The allocation to TAZs completes the cycle, for this data serves as the basis for trip generation for travel demand activities in the future year.

Before concluding this discussion, a few points are worthy of clarification. The sub-regional component of the process is ultimately the most difficult to perform, because it is the element with the least amount of automation and the largest number of zones. It represents a significant link in the chain of the land use and transportation modeling connection. On the latter subject, there are really two levels of potential integration between the land use and transportation modeling environments.

The first level should be considered as process integration. This involves developing compatible levels of geography and inputs and outputs that either process will accept and use. It also considers process flow so that inputs and outputs are created and delivered in proper sequence. The existing MAG land use and transportation protocols are already developed to accommodate this type of integration. Work needs to be performed on the sub-regional allocation model and software interfaces between the other elements to achieve complete process integration, but in general the geography and input and output are compatible. This is not typical of many MPOs or other planning organizations that deal with land use and transportation modeling.

The second notion of integration involves iterating the two components simultaneously until a satisfactory solution is derived. The most notable examples of this type of work involve the integration of DRAM/EMPAL with TRANPLAN at the Southern California Association of Governments (SCAG) and the integration of EMME/2 and DRAM/EMPAL at METRO in Portland, Oregon. In both of these cases, zone to zone travel times are imported from the travel demand models to DRAM/EMPAL. DRAM/EMPAL then produces a household and employment forecast, which is fed back into the travel demand model and new travel times are generated. These travel times are fed back to DRAM/EMPAL which produces another forecast and so on. The process is allowed to continue for a single forecast year until the percent change in the allocation of employment and households is minimized and the entire process reaches closure.

This is a simplistic rendition of the actual process, but the basic concepts are the same. As of this writing, the Portland process utilizes a simplified travel demand network with 100 TAZs, which corresponds to the RAZs utilized by DRAM/EMPAL. SCAG, on the other hand, takes the regional output from DRAM/EMPAL and allocates the change to a more detailed TAZ system before each iteration. The point of this discussion is to illustrate the different levels of integration that can be achieved between land use and transportation modeling, and to illustrate the relative importance of a sound sub-regional allocation process.

One final note, the integration of the two processes at SCAG and METRO are relatively new efforts, and the implications of this approach are not yet clear. In fact, in a recent paper describing these efforts, Steve Putnam laments that, "We need to examine these results to determine if they are indeed the better forecasts that we expect to have obtained."[1]

Data Requirements

There are two general categories of data that are required to effectively utilize the process mentioned above. The first group includes data that are required to calibrate and run DRAM/EMPAL. The second category includes data which must be allocated to the TAZ level to run the travel demand modeling process and support other information needs for the MAG.

---

[1] "Integrated Transportation and Land Use Models: An Interview of Progress with DRAM and EMPAL. With Suggestions for Further Research." Putnam, Stephen H., Transportation Research Board, Washington D.C. January 1994.

In order to correctly calibrate DRAM/EMPAL, the data items listed below must be aggregated to the RAZ level for the current year. Employment by type for five years previous to the year is also required. These items include:

Land Use

- Total area
- Undevelopable area
- Developed residential area
- Undeveloped residential area
- Developed non-residential area (and by type)
  - Retail
  - Office
  - Industrial
  - Public Administration
  - Other
- Undeveloped non-residential area

Socioeconomic

- Household income by quintile
- Population in households
- Population in group quarters
- Total employment
- Employment by type
  - Retail
  - Office
  - Industrial
  - Public Administration
  - Other

It should be noted that many of these items are summary variables which represent combinations of other data items that must be maintained at a level of geography lower then the RAZ. To produce future year forecasts, additional data must be supplied to DRAM/EMPAL. These items include:

- RAZ to RAZ peak hour travel times (for the current and each future year)
- Future year county level projections of employment by sector
- Future year county level projections of population

In addition, a number of other parameters have to be input into the process to temper the allocation. Examples of these include land absorption rates, employment and household attraction parameters, unemployment rates and household to population conversion assumptions.

As noted earlier, the data required to support this process is quite extensive. The implications are that a great deal of land use and socioeconomic data has to be maintained at lower levels of geography. This is good in many respects for the process of sub-regional allocation. Since much of the base data utilized to generate these summary variables is maintained at levels of geography equal to or lower than the TAZ. MAG already has a considerable amount of suitable data upon which to build a more robust process. This issue is discussed in greater detail below.

The variables that must be *forecasted* to support the travel demand modeling process and other MAG planning needs include:

- Total resident population
- Resident population in households
- Resident population in group quarters
- Transient population
- Seasonal population
- Total resident housing units
- Occupied resident housing units
- Total employment
- Retail employment
- Office employment
- Industrial employment
- Public employment
- Other employment
- Average household income
- Developed residential area
- Undeveloped residential area
- Developed employment area
- Undeveloped employment area
- Undevelopable area
- Total area

These items are forecasted to the TAZ level out to the horizon year of the travel demand modeling process (which has been 25-30 years from the current year) and even further at the RAZ level for other purposes. Most of these items are produced by, or derived from, DRAM/EMPAL RAZ forecasts.

It is the process of forecasting these items (and possibly others) to the TAZ level that the sub-regional allocation procedure must address.

Data Collection and Improvements

MAG has invested a great deal of time over the past several years to improve data and methods utilized in the planning process. Efforts to improve the collection and forecasting of socioeconomic variables were undertaken in 1990 and more recently in 1993. The 1993 endeavor, which was aimed at utilizing census products and the development of key databases to support DRAM/EMPAL, resulted in several new or improved products. These include:

- Better estimates of existing special population groups (resident, seasonal, transient and group quarters).

- The development of an employment database which includes existing data on all employees in Maricopa County with 50 or more employees.

- Geocoding and updating a parcel level database, which contains close to a million records.

- An improved TAZ structure which corresponds more readily with Census geography.

- The development of a database which profiles large scale planned and proposed development.

- Improved methods for forecasting population subdivisions at the county level.

- General Plan Land Use updates.

- Developing improved linkages between land use types and employment sectors.

A notable feature of this effort was the attention paid toward effectively linking data with the GIS environment. This has resulted in a situation where almost any data item can be related to any other data item for equivalent or alternative levels of geography.

MAG has also made important strides to improve the travel demand modeling process. Recent efforts have been completed, or will be completed soon, to improve several components of the models including:

- External trips
- Mode choice
- Commercial vehicle trips
- Special generators
- Parking costs

Most of these projects included extensive data collection efforts, including household, on-board or roadside surveys, traffic counts, travel time runs and vehicle occupancy surveys.

The models were also recently converted from the Urban Transportation Planning System (UTPS) software environment to EMME/2. Additional work will begin soon to add more flexibility and sensitivity to various model components. The work that has been completed in the past has greatly improved MAGs capability to provide effective planning products. It also suggests that improvements to the process have been constantly oriented toward process improvement and integration. The existing data and process structures for land use and travel demand forecasting certainly reflect this intention.

1.2 Sub-Regional Allocation

Based on literature review and discussions with contacts in other agencies, it is clear that a considerable amount of emphasis has been placed on the county to regional allocation (RAZ equivalent), but very little attention has been paid to allocations to lower levels of geography. One of the reasons for this may be that regional allocations satisfy most of the data requirements imposed by the planning process in most areas. On the other hand, most of these same agencies have to make estimates to the TAZ level to support travel demand modeling efforts. It may also be that recent efforts to integrate land use and transportation models put an immediate emphasis on solving the county to regional interface, and sub-regional allocation has yet to be addressed in this framework.

In either case, it is clear that many agencies have methods for performing sub-regional allocation, but typically these processes require manual assessments and intervention and have not been developed to the level of sophistication that one finds in the regional counterparts. However, there do not appear to be major obstacles to doing so.

The initial part of the discussion below focuses on general approaches to sub-regional allocation utilized by other areas. This information was compiled from a series of phone conversations with key people at the following agencies.

- SCAG (Southern California Association of Governments)
- ABAG (Association of Bay Area Governments)
- NCTCOG (North Central Texas Council of Governments)
- HGAC (Houston Galveston Area Council)
- SANDAG (San Diego Association of Governments)
- City of Seattle
- METRO (Portland Metropolitan Service District)
- SRP (Salt River Project)
- MRGCOG (Middle Rio Grande Council of Governments-Albuquerque, NM)
- SNL (Sandia National laboratories)

Subsequent to this discussion, a brief overview of some key issues related to sub-regional allocation is presented.

Survey of Existing Practices

The phone survey upon which these remarks are based elicited different levels of interest and information from the parties contacted. These agencies were contacted because almost all use DRAM/EMPAL to perform regional allocations. The respondents were asked specific questions related to the degree to which the process was automated, the use of GIS in the allocation procedure, and key variables upon which the allocation was based. An effort was also made to collect information on differences between regional and sub-regional procedures. A brief description of results of these conversations is presented below.

SCAG

SCAG utilizes a Small Area Model (SAM) to allocate regional variables to TAZ equivalents. Forecast variables include households by type and income, licensed drivers, vehicles and employment. The sub-regional allocation is based on existing distributions, vacant land and assumed densities and the regional forecasts from DRAM/EMPAL. This allocation is performed by distributing growth to the existing condition or prior forecast year, based on relative levels of opportunity derived from regression equations. The GIS is utilized to store and maintain data used in the process, however, the SAM operates external to the GIS environment.

ABAG

ABAG also maintains a SAM which allocates population and employment to the census tract level. This allocation is similar to that utilized by SCAG, although it does consider travel times as a measure of accessibility. The actual allocation is also influenced by local land use policies and planned and proposed development. GIS is not utilized directly in the process, but it is used to create maps and other products.

NCTCOG

NCTCOG utilizes DRAM/EMPAL to allocate population and employment to a district level. The process for further allocating variables to the TAZ level was described as a manual additive procedure that considered local knowledge, future zoning existing data and land ownership patterns. They are currently working to update several land use coverages in the GIS, to improve inputs for DRAM/EMPAL and the sub-regional process.

HGAC

HGAC utilizes DRAM/EMPAL to allocate population and employment to 199 regional zones. These variables are further allocated to TAZ by pro-rating growth to the TAZ level based on available land, adjacent zone patterns, planned development, the HGAC comprehensive plan, and environmental constraints (soil and ground water contamination). They are currently in the process of evaluating how to bring zoning into this process. The allocation is performed manually based on these factors. GIS does not play a formidable role in this process, but they do have GIS capabilities and would like to integrate it with this process.

SANDAG

SANDAG also utilizes DRAM/EMPAL to forecast population and employment to the regional level. These variables are then allocated to the census tract level and once again to the census block level, utilizing a fortran based SAM. The allocation is based on accessibility, capacity for development, planned developments and existing distributions. SANDAG makes extensive use of the GIS to support this process (mainlining coverages and data), but the allocation is run separately.

City of Seattle

Seattle receives DRAM/EMPAL based regional forecasts of population and employment that are produced by the Puget Sound Regional Council. Allocations to TAZ are performed in a two step process. Seattle utilizes a small area model to perform the initial allocation utilizing development capacity, general plan land use assumptions, environmental constraints, existing and planned distributions and land value. The resulting estimates are then evaluated and manual redistributions are performed usually to compensate for redevelopment, which is formidable factor in this process. It is interesting to note that Seattle utilizes the ratio between a parcels land value and property value to identify candidate areas for redevelopment, and expect to more fully integrate this in the process. The GIS is not used directly in the allocation process, but is utilized extensively to provide inputs and analyze results.

METRO

METRO has been extensively involved with S.H. Putnam & Associates (and many others) to improve the land use and transportation components of their planning process. The travel demand models have been overhauled and integrated with DRAM/EMPAL (utilizing the feedback loop) at two levels. In one case, a sketch level transportation network (with 100 zones) is utilized to iterate directly with the regional outputs and inputs required by DRAM/EMPAL. No sub-regional allocation is required for this process. A more detailed representation of the transportation network can also be utilized. This network includes 1189 TAZs and requires sub-regional allocation of data produced DRAM/EMPAL at the 100 zone level. Unfortunately the more detailed process is still under development and information about the sub-regional allocation could not be ascertained through the phone survey. There is a secondary effort underway to produce a policy based, land use allocation (or aggregation) which may be of some interest. In this process county control totals for population and employment are "allocated" to quarter acre grids in a GIS environment. The allocation is based on parcel level data, comprehensive plans and three categories of land use. These categories include developed land, re-developable land and vacant land. The allocation is further constrained by environmental considerations and policy based area density profiles. These density profiles relate specifically to urban centers, transportation corridors, residential and commercial areas. The allocation is performed at the cell level (considering all these factors) utilizing grid in ARC/INFO. The results of this allocation were aggregated to the 100 zone level, so a comparison could be made with the outputs from DRAM/EMPAL. Apparently there were some differences, but overall the results were close. It was noted that the aggregation technique was developed to test land use policies and was therefore oriented toward simulating "what we would like to happen" rather than following the more predictive approach utilized by DRAM/EMPAL.

SRP

SRP is the municipal utility for the greater Phoenix area, and is charged with providing their service area with water and electricity. As a result their long range planning process has to accommodate the allocation of growth for a wide variety of internal uses, including the evaluation of future levels of electric power and water usage. Therefore, the sub-regional allocation procedure is significantly different than the examples provided above. Basically, 40 acre tracts of land are represented as grid cells in the GIS. In most cases, the 40 acre cells represent relatively homogeneous land uses. An application has been developed which allows analysts to develop "preference maps" for several different types of land use. A preference map is developed by weighting several factors relative to the land use type. In some cases up to 13 factors will be considered in the development of a preference map. For example, a preference map for light industrial use might consider adjacent 40 acre cells, and apply negative ratings if the areas are primarily residential. Alternatively, 40 acre zones in proximity to freeways, rail and warehousing would receive favorable weights. Preference maps are generated for the entire service area for 30 to 40 different land uses. In addition, this process is completed for the 20 year horizon and at five year intervals in between. The preference maps for each five year period are then "loaded" into a land use allocation program, which allocates acres of growth by land use type. It was not all together clear exactly how this process works but apparently land use types compete against other land use types for vacant land (based on the weights for each cell by type). The end results of this application is the distribution of acres by land use types to the 40 acre zones.

MRGCOG

The MRGCOG employs several different techniques to allocate various socioeconomic characteristics to the regional level, and then again to TAZ. Neither of the processes are fully automated, although algorithms have been developed to allocate several variables. County to regional allocations are based on several general considerations including the availability of bulk land, planned development, building permit trends, land use, infrastructure and existing distributions. Population is allocated to the regional and TAZ level based on these considerations and an attractiveness factor which is a function of capacity, the percentage of residential land that is developed, existing median housing prices and travel time to the Central Business District (CBD). Employment is allocated manually in both cases and considers zoning vacant buildings and the number of employees per acre and by square foot for alternative employment sectors, in addition to the regional factors mentioned above. MRGCOG utilizes the GIS extensively in this process, to maintain and manage data and to assess results.

SNL

As part of a recent research and development program, Sandia National Laboratories, in cooperation with BA and Resource Systems Group, developed a prototype integrated modeling environment which performs trip generation, EMME/2 based travel demand modeling, land use allocations and generates mobile source emission profiles. These processes are all launched from the ARC/INFO GIS environment from a graphical user interface (GUI) so that background operations are transparent to the user. Results from all processes can then be accessed from the GUI. The process does not incorporate a feedback loop per SCAG or METRO, but does include a land use model (called *ithink*) which allocates dwelling units and employment to a regional level. These characteristics are then allocated to the TAZ utilizing a very basic sub-regional allocation model in ARC/INFO. *Ithink* is a modified systems process model which utilizes existing land use data (housing stocks and land use), construction, capacity and growth rates, and impedances from EMME/2 to predict the distribution of growth. The effects of all variables in all regions are calculated to update each variable. Sub-regional allocation utilizes the regional outputs from *ithink* and allocates growth utilizing an additive function which considers existing distributions of population and employment only.

The major objective of this initial effort was to demonstrate that all of these components could be effectively chained and managed from the ARC/INFO environment. From the GUI the user can launch the entire process (after specifying key inputs) and the system will generate data for as many as 10 time steps without any additional user input or intervention. The entire system was designed with modular elements, in recognition of the relative weakness of some components (namely the sub-regional allocation process).

Discussion

There are several salient points that can be made about the sub-regional allocation procedures currently in use. First, it was very difficult to obtain much detail about these processes over the phone. These conversations rarely got to a level of detail that allowed for specific questions related to the use of the Census Transportation Planning Package (CTPP) in the allocation process and the treatment of special population groups or usage of floor area ratios. Follow up calls could be made to address these issues, provided these individuals are willing to take the time to provide information at a level of detail sufficient for use in this study. In most cases the calls were directed to individuals involved in the process, but only a general representation of the process was obtained due to time constraints. In many cases answers to questions were vague, particularly when asked about levels of automation and variables considered in the process. However, some general conclusions can be drawn from the interviews.

Most agencies use a tiered process, in which regional growth is allocated to the base year condition, or prior forecast year in an *additive* fashion. It is also clear that the primary reason for performing sub-regional allocation is to provide data at a level suitable for the travel demand modeling processes used by these agencies. A composite assessment suggests that there are many different elements that are considered in the sub-regional allocation procedures utilized by these agencies.

These include attributes associated with the allocation surface such as land use, zoning, land ownership, adjacency, value, and constraints. Accessibility is addressed at less explicit levels in some cases and directly in others (normally in the form of travel times or impedances). Because sub-regional allocation is normally associated with TAZs, incorporating zone to zone travel times as a measure of accessibility is also mechanically feasible.

Since most of these processes output additive measures, land consumption needs to be addressed through each time step. In some cases this is handled through predefined capacity measures (a TAZ can accept growth until the maximum is reached), and in other cases it is handled in terms of area either directly (acres consumed) or indirectly (via converting households, population, employment to some measurement of area via assumed densities).

Redevelopment was an issue that did not surface as a major component, except in areas that were close to build-out (City of Seattle). In this case a specific variable was generated to identify candidate parcels for redevelopment. Even in this instance, the variable was not incorporated as a direct element of sub-regional allocation. In other agencies, redevelopment was handled less explicitly, either through redevelopment plans or as an element of the region's comprehensive plans.

Although very little information was obtained about specific procedures in cases where the allocation was not automated, general statements tended to indicate that a combination of techniques were employed, including Delphi allocations, trends analysis and shift share.

In all, there are several principles that most of these sub-regional allocation processes recognize. First, there is the notion of capacity, or the upper limit of development that can be considered for a particular zone, based on assumed densities and land available for use. Second, growth is allocated based on relative measures of attractiveness generally expressed as a gradient reflecting the least amount of access or opportunity to the greatest amount of access or opportunity within a particular region. These attractiveness measures are based on the data available to calculate them, and the attributes which must be allocated. Models utilized in the process ultimately represent a mechanical set of rules, under which the allocation is played out. The results are a reflection of the rules employed and the treatment of variables related to the surface and data elements which are allocated. Finally, given the amount of data which must be maintained and managed in this process, GIS was recognized as a critical element either as a direct component of the process, or simply as an environment for the creation, management and review of data utilized in the process.

A cursory literature search was also conducted as part of this process. Most of the literature evaluated for this report (related to allocation) was oriented toward the regional level process. One particularly comprehensive source on existing land use models and activities is "Volume 1 Modeling Practices" of the 1000 Friends of Oregon series. This report profiles existing land use models, and provides general and specific distinctions between three different model groups. The first group (which includes DRAM/EMPAL) is differentiated by the almost exclusive use of accessibility to allocate socioeconomic variables. The second group (which includes MEPLAN, MEP and TRANUS) are recognized for their emphasis on competition and land values as a means for allocation. The last group (which includes TOPAZ and TOPMET) are noted for their orientation toward urban form and design and the lack of predictive elements. These models produce allocations based on policy or objectives.

It is beyond the scope of this report to discuss these models in greater detail. It is important to recognize that alternative orientations could be considered in the design of a sub-regional process. This may be particularly relevant when considering agency demands to evaluate alternative urban forms or divergent regional land use directives which may require allocations that do not follow predictive patterns. There is also a parallel notion that sub-regional activities should be compatible with the larger processes, not only in terms of their treatment of inputs and outputs, but also accommodating data flows in two directions. This is an important aspect, if the sub-regional allocation results in data that must be aggregated back to the regional level for use in subsequent time steps.

1.3 MAG Data Sources

As noted previously, MAG has devoted a considerable amount of time and effort to the development and maintenance of these data sources, and has also insured that most data files can be linked to geography and can therefore be related to the GIS environment.

Generally speaking, these data fall into the following categories:

- Land use inventories which include existing and planned land use, parcel level information, planned and proposed developments and land ownership characteristics.

- Socioeconomic data profiles which include information on existing and forecasted population, household, employment, and income characteristics.

- Geographic boundaries/features (upon which most of this data can be related) including coverages which represent bookmap, land use, TAZ, RAZ, MPA, Census and township boundaries, and coverages which profile street networks, elevation contours and air quality non-attainment boundaries.

In addition to these sources, MAG is in the process of uploading tables from the CTPP. These tables can be related to MAG 1990 TAZ geography and include information on population and household characteristics, place of work profiles and journey to work characteristics by zone pairs.

A review of the MAG data sources will provide a better appreciation of the types of data available and the relationships between data sources that can be called upon in the process of improving the MAG sub-regional allocation methodologies.

2
Prepare Networks, Databases, and EMME/2 Interfaces

Travel forecasting networks play a critical role in a number of MAG's functional responsibilities. They supply measures of link level VMT (vehicle-miles of travel) that figure prominently in emissions estimates necessary for air quality conformity analyses required by the CAAA90. They also supply measures of accessibility that drive the land use forecasts generated by the DRAM/EMPAL model chain. In the past several years, MAG has made substantial improvements to its travel forecasting models, and has modified the software and hardware environment in which they operate.

In this element of the GIS analysis and data enhancement program, the forecasting networks in use at MAG were geocorrected to conform with the actual underlying geography they represent. Formerly, MAG travel networks were digitized according to a MAG-unique coordinate system. This approach has been sufficient to support graphics and plots associated with travel model output. The emergence of GIS at MAG, however, suggested a need to correct the underlying coordinate system by which the travel networks were digitized. The advantages associated with this are:

- Travel networks and their associated attributes can be plotted with background information (e.g., street name, TAZ boundaries, census boundaries, or landmark features) coming from other GIS sources.

- Key attributes associated with travel networks (especially link length) can be automatically computed, not scaled, by the GIS or the forecasting package. This assures greater accuracy and consistency.

- Appropriately geocorrected travel networks is a step toward integration of travel forecasting databanks and GIS, thereby making the visualization and analytical power of GIS available for travel forecast data.

2.1 Task Objectives

The specific objectives of this task were as follows:

- Geocorrect an initial set of five MAG networks, including the 1991, 1995, 2001, 2005, and 2020 networks. Later, an additional four networks were also geocorrected, including the 1993, 1995, 2005, and 2015 networks. The networks were geocorrected to Arizona Central State Plane coordinates using the street centerline coverage for Maricopa County as a background coverage.

- Nodes were renumbered following a general geographic sequence. New node numbers were then propogated across all of the four active networks so that nodes appearing in all networks would have the same identifier. Also, turn penalty and transit line description files were modified to reflect new node identifiers. Lookup (equivalency) tables which will enable the translation of old node numbers to their new equivalents were constructed.

- AMLs to accomplish these objectives, as well as to import and export forecasting networks between ARC/INFO and EMME/2, were developed.

- Any changes required by MAG travel demand forecasting programs were made. In fact, by preserving TAZ numbers, the only changes necessary involved computations that were predicated on coordinates (e.g., distance computations based on TAZ coordinates). These were installed by MAG staff.

- Network modifications necessary as a result of a new TAZ structure were deferred until conversion to the new TAZ structure (covered in Chapter 3) is imminent .

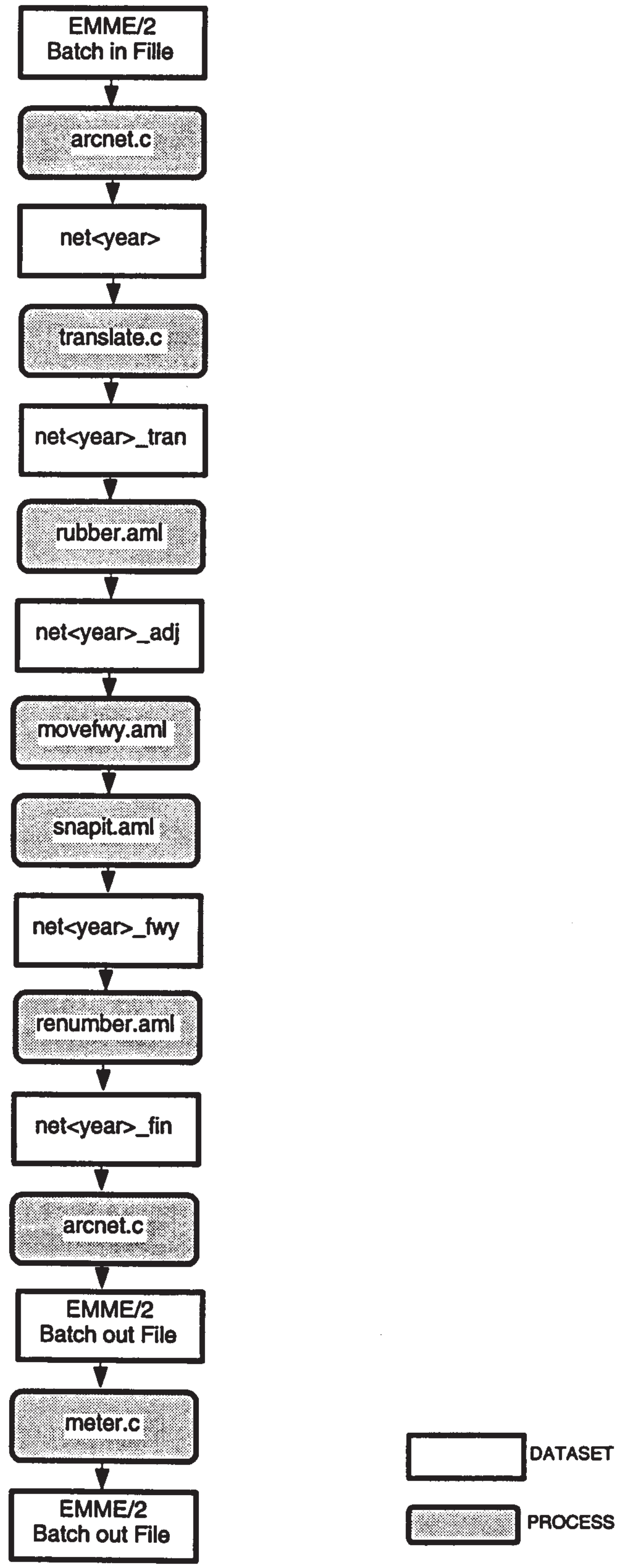
2.2 Methodology

Almost all of the work associated with revisions to MAG networks was performed in a GIS context. The process involved a number of elements, as follows (also, see Figure 2-1):

- *EMME/2 Import/Export:* A mechanism to import and export travel forecasting networks between the ARC/INFO and EMME/2 database environments was developed. This software, a combination of ARC/INFO AML and a custom-written C program, was used to build a GIS coverage of the forecasting networks from EMME/2 *batchin* files. The coverage produced from the original networks were named *<netyear>* (as in *net1993*, *net1995*, etc.)

- *Network Scaling/Translation:* Rubber-sheeting networks to match an underlying background coverage or basemap works best when both appear onscreen at the same time. For this to be the case, they both must share the same approximate

Figure 2-1
Network Geocorrection Process



coordinate space. A custom-written C program was developed to translate and scale networks which were originally coded according to arbitrary map-inch coordinates into Arizona Central State Plane coordinate space (NAD27). This space is the same as the street centerline basemap (that used fix positions). The coverage produced after this process was *<netyear_tran>* (as in *net1993_tran*).

- *Rubber Sheeting:* Rubber-sheeting in a GIS environment involves the coding of directed *links* (in ARC/INFO terminology). The ARC/INFO notion of a *link* is not to be confused with the traditional usage adopted by transportation planners, as in *roadway link*. ARC/INFO *links* are special directed arcs which indicate to the GIS how features are to be moved (from the old location to the new location). The entire map is then rubber-sheeted accordingly. A menu-driven AML to facilitate the coding of links for each individual network was developed for the technician's use. The coverages produced from this process were called *<netyear_adj>* (as in *net1993_adj*, etc.)

- *Freeway Adjustments:* The rubber-sheeting process ignored freeway interchanges, because of their complexity and unsatisfactory results during trials (freeway alignments were warped). A menu-driven AML was developed to facilitate editing of freeway links (by literally moving nodes, dragging freeway and ramp links along with them) manually. In addition, since the manual editing process was so time consuming, another special AML was developed to migrate coordinate changes to freeways made in one network to other networks. The coverages produced from this process were called *<netyear>_fwy* (as in *net1993_fwy*, etc.)

- *Node Renumbering:* Nodes (that is, non-centroid nodes) were sorted geographically (by their *x,y coordinates*) and then renumbered. This has the effect of renumbering nodes laterally, west to east. In order to preserve node numbers from one network to the next, a somewhat complicated series of operations were designed to propogate node numbers from one network to the next. The complication, as will be described later, has to do with how nodes in one network should be matched to their comparable member in another network that was already renumbered. The coverage produced from this process was named *<netyear>_fin* (as in *net1993_fin*, etc.).

- *Ramp Meter Links:* Freeway ramp links were initially adjusted to match the actual shape and configuration of the actual physical feature they represented. This involved using nodes associated with ramp links which represented ramp meters as a *shape point*. Later, MAG expressed a preference that meter links should have 0-length, since they did not really represent an actual physical feature. Consequently, a custom-written C-program was developed to locate all of these links in a network and adjust their coordinates accordingly. The coverages produced from this process were called *<netyear>_meter* (as in *net1993_meter*, etc.)

- *Turn Penalties and Transit Lines:* Two other datasets associated with a travel forecasting network are closely associated with node identifiers. Locations in the network where turn penalties are in effect are identified by node number. Also, transit routes are described in terms of the path (sequence) of nodes which are traversed. Node numbers appearing in these files had to be modified to reflect the new numbering scheme.

Once networks were geocorrected and renumbered, the import/export facility was used to create EMME/2 batchin files from the ARC/INFO coverages.

Each of the components of the geocorrection process will now be described in greater detail.

2.3 EMME/2 Import/Export Facility

A facility for moving network databases between ARC/INFO and EMME/2 was developed. The application was written principally in AML, and is menu-driven. However, because the internal data structures representing forecasting networks used by ARC/INFO and EMME/2 are so different, customized C programs were developed to extract and reformat network information.

The fundamental difference between ARC/INFO and EMME/2 network representation involves the way geographical features are represented in data structures. In EMME/2, the principal feature in a network is the *node*. Coordinates, for example, are associated with nodes; link locations are known through pointers to the nodes. That is, the *key* to a link is the *inode, jnode pair* which defines the link. The location of the link can be determined by retrieving the coordinates associated with the inode and the jnode.

The principal feature in ARC/INFO coverages, however, is the *link* (the *arc*, in ARC/INFO terminology) — nodes are purely incidental and are defined when arcs intersect. The geography of a network is determined by the pair of coordinates that define the link location.

This means that an import/export operation between the two environments has to contend with *more* than just reformatting files. When moving data from EMME/2 to ARC/INFO, coordinates for each link have to extracted from EMME/2 and written to a flat-ASCII file. This involves (1) building a node table with node coordinates, (2) scanning each link in the network for the inode,jnode pair defining the link, (3) searching the node table for the inode and jnode, (4) retrieving the coordinates associated with them, and (5) writing the coordinates for the link to a file. Since ARC/INFO keeps geographical information (coordinates for features) in separate files from the attribute information, *other* data associated with EMME/2 links (e.g., modes, facility type, etc.) must be imported to ARC/INFO separately and then *joined* to the coverage.

The export operation designed to move data into an EMME/2 batchfile from ARC/INFO involves similar operations run in *reverse*. ARC/INFO coordinates associated with links must be moved to the nodes associated with the end-points before being written to a file. And since native ARC/INFO provides *two different commands* for exporting geographical information and attribute information, a special program must be written to read both export files and join them back together in order to create an EMME/2 file.

Network Database Structure

The EMME/2 network database structure, involving node and link records, is described in Figure 2-2. *Node records* (1) identify node numbers, (2) report the x,y coordinate position of the node, and (3) report attribute information associated with the node. Attribute information includes a flag to designate nodes which are zone centroids. Also, up to 3 floating point user-defined fields are provided, although this feature is not used by MAG. *Link records* only identify the beginning and ending node numbers (the *inode* and *jnode*). All of the rest of the information on a link record is attribute information. Internally, EMME/2 can associate link and node information by way of the node number — this is the key to each of the records.

Figure 2-2
EMME/2 Database Structure

Record Type	Field Descriptions
Node	flag to describe disposition of the node (e.g., "a" for append) flag to distinguish centroids from nodes node identifier x coordinate y coordinate up to 3 user-defined floating point fields
Link	flag to describe disposition of the link inode identifier jnode identifier link length modes permitted on the link link type classification number of lanes on the link volume-delay function up to 3 user-defined floating point fields

This same information can be easily represented in ARC/INFO also, although the file structure is different. There are, in fact, three files involved:

- The *arc* file for a coverage is a *binary file* which stores the *coordinate pair*[1] associated with each link.

---

[1] In fact, the arc file can store more than a single coordinate pair describing an arc. ARC/INFO allows shape points, so the arc file can hold an entire sequence of coordinates describing a link (in fact, arc features in ARC/INFO are polylines). Data coming from EMME/2, however, always involve single line segments and one coordinate pair.

- The *nat* file is an info database file which can store node numbers and node attribute information. On request, ARC/INFO can also generate coordinate information associated with nodes. The nat file is not normally built for ARC/INFO coverages, but can be generated at the users request.

- The *aat* file is an info database file which can store attributes associated with arcs (that is, network links).

One-Way and Two-Way Link Representation

Network representation in older travel forecasting packages (e.g., UTPS) considered roadway links to be single two-way links. EMME/2, however, represents networks as directed 1-way line segments, so there are two links commonly associated with each roadway segment (each going in opposite directions). The ARC/INFO forecasting network is represented in the same way — although this is somewhat contrary to ARC/INFO's usual view of the world:

- ARC/INFO doesn't typically expect a pair of nodes to be connected by two links — one running in one direction and the other running in the opposite direction.

- Since both links share the same coordinates, they are superimposed on the screen. This can cause problems for ARC/INFO users attempting to select individual links.

Still, even though the data model for network representation in ARC/INFO did not conform with ARC/INFO's typical expectations, no problems were encounter with it.

The Conversion Process

A user-friendly AML application was implemented to automate the transfer of networks between the ARC/INFO and EMME/2 environments. This application, called ***arcnet*** [2] provides the user with a menu to direct importing and exporting of networks. The AML relies heavily on two C-programs (***arcm2.c*** and ***m2arc.c***) to reformat EMME/2 batchin files into a form more easily accommodated by internal ARC/INFO commands. The features of the application allow users to name user-defined fields (allowing for something more expressive than UL1, UL2, and UL3 as in EMME/2). Note also that ARC/ INFO will allow an unlimited number of attributes to be associated with links, while EMME/2 is severely restricted. Therefore, with an import/export facility GIS can be used to support expansive network databases in EMME/2. Another feature of the application is to flag loadlinks (zone centroid connectors) so that they may be more easily selected when working with the coverage (for example, to draw them in a different color).

[2] In fact, arcnet also provides a mechanism for importing/exporting MinUTP networks as well. This feature was added in connection with another project.

The details of the network representation in ARC/INFO is shown in Figure 2-3.

Figure 2-3
ARC/INFO Database Structure for Networks

File	Field Name	Width	Data Type	Description
AAT	<COVER>#	4	BinInt	Internal Link ID maintained by ARC/INFO
	<COVER>-ID	4	BinInt	Internal User Link ID maintained by ARC/INFO
	FNODE#	4	BinInt	Pointer to the *from node* in the NAT, maintained by ARC/INFO
	TNODE#	4	BinInt	Pointer to the *to node* in the NAT, maintained by ARC/INFO
	LENGTH	4	BinReal	Link length maintained by ARC/INFO
	INODE	6	Int	Network inode from EMME/2
	JNODE	6	Int	Network jnode from EMME/2
	LOADLINK	1	Char	"*" flags loadlinks
	LEN	6.2	Real	EMME/2's record of link length
	MODE	12	Char	String identifying eligible modes
	TYPE	3	Int	Link Type Code
	LANES	4.1	Real	Number of lanes
	VOLDEL	3	Int	Volume-delay function number
	User Fields	12.4	Real	User Fields
NAT	<COVER>#	4	BinInt	Internal Node ID maintained by ARC/INFO
	ARC#	4	BinInt	Pointer to ARC
	NODEID	6	Int	EMME/2 Node Number
	CENTROID	1	Char	"*" flags centroids

Problems Encountered

Several problems were encountered when converting MAG networks into ARC/INFO coverages:

- MAG networks originally contained disconnected nodes, existing by themselves, but not connected to any links. (This is probably done to facilitate adding links for future year networks.) ARC/INFO can not tolerate this type of topology. ARC/INFO data structures define nodes to be nothing more than the terminals for links — without a link there can be no node.

  There are potential work-arounds for future consideration involving a process that would dump isolated nodes into an ARC/INFO point coverage. This, however, would not have helped us in this project since there would be no way to identify how these disconnected nodes should be rubber-sheeted (since they do not relate in any apparent way to an underlying geographic feature in the background street system coverage).

- MAG networks contained a number of 0-length links. A 0-length link occurs when both the inode and the jnode for the link occupy the same position, exactly. There might be reason for networks to be coded this way (for example, to represent non-physical features in a network such as ramp meters), however it probably is not good practice — users would have problems viewing and selecting nodes onscreen. In any event, while 0-length links are tolerated in forecasting software such as MinUTP and EMME/2, they can not reside in ARC/INFO — these links are rejected. Consequently, the conversion process developed for importing EMME/2 networks was revised to look for these instances. When found, one of the nodes had to be moved (an arbitrarily small distance) to give the link some length.

Installation

The *arcnet* application was installed on the MAG computer system in the directory */magtpo/m2*.[3] Listings of program code associated with the application can be obtained from there. To run the application, merely move to that directory and type:

&run arcnet

ARC/INFO network coverages are expected to be in the */magtpo/m2* directory. EMME/2 batchin files are expected to be in the */magtpo/m2/data* subdirectory. (This is true for both imports and exports from ARC/INFO). From there, files and coverages can be moved elsewhere on the system, wherever desired. The application is otherwise menu-driven, and usage of the program is therefore readily apparent.

2.4 Network Scaling and Translation

Rubber-sheeting networks to match a background street coverage involves coding directed line segments (called *links* in ARC/INFO terminology), indicating how nodes in the network should be moved. The collection of ARC/INFO *links* shows how the coverage is to be warped, or rubber-sheeted. The editing process involved with coding ARC/INFO *links* works best when both coverages, the network to be rubber-sheeted and the background coverage to be matched, appear on the screen at the same time. This requires that coordinates associated with input network coverage be translated from the MAG-unique map-inches coordinate system to Arizona Central State Plane (ACSP).

This was accomplished through a custom-written C program (***translate.c***) which reads coordinates from an ARC/INFO *generate* file (a text file of coordinates exported from ARC/INFO), converts them to (approximate) ACSP coordinates, and saves the changes back to an ARC/INFO *generate* file. The file is then reloaded into ARC/INFO and attributes

---

[3] Since MAGTPO's licenses for ARC/INFO and EMME/2 are operating on different machines on the network, the selection was a directory that is visible from both. The MAGTPO directory on redmtn is the same as the drive2 directory on magtpo.

associated with nodes and links are rejoined to complete the network coverage. Coordinates are changed from map-inches to Arizona Central State Plane through a simple linear translation which both translates (moves points a fixed distance) and rescales coordinates.

The general equations for a simple linear coordinate translation to convert map-inch coordinates to ACSP coordinates are as follows:

$$x_{ACSP} = a_x + b_x x_{inches}$$

$$y_{ACSP} = a_y + b_y y_{inches}$$

The constant coefficients a_x, b_x, a_y, and b_y applicable must be *derived.* This was done by finding two locations common to both maps, preferably diagonally opposite each other (e.g., in the far southwest of the region and the far northeast of the region) so as to maximize the variation in the *x* and *y* directions. With the coordinates in both map-inches and ACSP known, the constant coefficients can be solved by *simultaneous equations*, resulting in the following linear transformation:

$$x_{ACSP} = 270{,}091 + 5280x_{inches}$$

$$y_{ACSP} = 749{,}761.6 + 5280y_{inches}$$

That is, it appears that the original coordinates associated with the EMME/2 networks were coded from a regional map with a scale of 1 inch to the mile.

The program merely asks for the name of the input ARC/INFO *generate* file and where the output should be written. The output file can then be read by ARC/INFO (the ***generate*** command) to create a new coverage in ASCP coordinates. The process of joining attribute information from the original network coverage to the new translated network coverage requires able knowledge of ARC/INFO.

The coverage created by the scaling/translation process are shown in Figure 2-4.

Figure 2-4
Network Coverages Created During Network Scaling/Translation

Subject	Original Cover		Translated Cover
Network Scaling	net93	»	net93_tran
	net95	»	net95_tran
	net05	»	net05_tran
	net15	»	net15_tran

2.5 Rubber Sheeting

Geocorrection can start once the target network shares the same coordinate space as the background street centerline coverage. Rubber-sheeting is a native ARC/INFO capability which requires users to establish a coverage of directed arcs (called *links* in ARC/INFO terminology) which show source and destination points for how the network is to be moved. The technique was to code *links* throughout the network from all nodes in the network to comparable intersections in the background street centerline coverage. Also, the links were snapped to nodes in the street centerline coverage, thereby achieving as close a fit as possible.

Coding links from the interactive command interpreter in arcedit can be a laborious, time-consuming process. Consequently, an application (***rubber.aml***) was developed which provides the user with menu-driven buttons and commands to assist in the process. The application offers the user many features, including the capability to declare background and snapping covers, the ability to add, modify, and delete *links*, and the ability to view the results according to any window on the network. It is, for all practical purposes, an onscreen menu-driven editor for coding links.

ARC/INFO *links* can also be saved for future use on other future-year networks. Therefore, the actual manual coding that was required for future year networks was more efficient because it could build on previous work performed for prior year networks.

Once rubber-sheeting was performed, it was also attempted snapping nodes in the networks to comparable nodes in the street centerline coverage. This proved to be a risky activity, because ARC/INFO will start snapping all nodes within whatever tolerance distance is established when snapping functions are activated — the resulting network loses its original topology when nodes are snapped together (*in other words, the network loses nodes*). Since MAG networks include numerous nodes which are located virtually on top of one another, it was found that node snapping generated unsatisfactory results.

The coverage created by the rubber sheeting process are shown in Figure 2-5.

Figure 2-5
Network Coverages Created by Rubber-Sheeting

Subject	Original Cover		Rubber-Sheeted Cover
Rubber Sheeting	net93_tran	»	net93_adj
	net95_tran	»	net95_adj
	net05_tran	»	net05_adj
	net15_tran	»	net15_adj

Installation

The rubber sheeting AML has been uploaded to MAG computers and is installed in the */magtpo/net* directory.

2.6 Freeway Adjustments

Initially, freeways and interchanges were left *as is*. Freeways in the original network appear to have been digitized to generate a *distinctive appearance* when viewed on a regional basis (that is, they were coded symbolically). In true geographic scale, however, this implied that median widths of freeways could be shown as wide as 500-feet. After examining several examples, MAG decided that a truer geographic representation of freeways and interchange complexes was preferred.

It was found that the rubber sheeting process did not work satisfactorily on freeways — it tended to warp the appearance of freeways which were actually following straight line alignments. This meant that freeways (along with interchange ramps and frontage roads associated with them) had to be manually edited to conform with the background coverage. In order to facilitate this process, an ARC/INFO application (***snapit.aml***) was developed. This application allows the user to move nodes in the target network to conform to the underlying location and shape of the street centerline coverage.

Alignments associated with future freeways, of course, could not be geocorrected with respect to the existing street centerline coverage. Without any background data to go by, it was matched the basic alignment and bearing followed by a freeway, connected interchanges to crossing arterial streets appropriately, and generally coded alignments for appearance.

The process of geocorrecting freeways and interchanges manually proved to be very time-consuming, even with the assistance of the ***snapit.aml*** application. In order to expedite the process, another application was developed to propogate the alignment changes made in one network to other future year networks. This application (***movefwy.aml***) relocates nodes associated with freeway links to the positions occupied by their peers in networks geocorrected previously.

The problem associated with this technique involves accurately matching freeway nodes appearing in two different networks. Since MAG networks for the various forecast years apparently have independent origins, node numbers are not consistent from one network to the next (they are usually, but not always). Therefore, node numbers can not be used to match identical nodes in two different networks. It was decided instead to match nodes on the basis of position under the presumption that two freeway nodes that were geocoded to the exact same location in two different networks represented the same thing, regardless of what their node number was. Therefore the process involved building a lookup table that related nodes in a *base network* that was already geocorrected with nodes in a *target network* that required geocorrection:

1. Extract the portion of the base network that represented freeways. This was easily done by selecting on roadway classes with volume-delay function numbers that end in 1, 7, or 8. The ***original networks***, coded in map-inch coordinates, were used as a basis for matching freeway nodes.

2. Through a computed relate, new coordinate positions for the base network were obtained and added to the look-up table.

3. The AML ***movefwy.aml*** was used to select freeway nodes in the target network and move them to new positions obtained from the look-up table.

This method significantly reduced the time to geocorrect freeways from about 8 person-days to 2 person-days. Unfortunately, MAGTPO network coding procedures offer no way to identify collector-distributor and frontage roads in the forecasting networks automatically (there are no roadway classes that distinguish them). Therefore, frontage roads still had to be edited manually.

The coverages created from the freeway editing process are shown in Figure 2-6.

Figure 2-6
Network Coverages Created after Freeway Editing

Subject	Original Cover		Freeway-Edited Cover
Rubber Sheeting	net93_adj	»	net93_fwy
	net95_adj	»	net95_fwy
	net05_adj	»	net05_fwy
	net15_adj	»	net15_fwy

Problems

A number of difficulties was encountered when attempting to edit network coverages created from native ARC/INFO rubber-sheeting commands (*arcedit* system crashes with fatal errors). Numerous discussions with ESRI failed to resolve the problem, which continues to have an unknown source. Exporting and rebuilding network coverages to *<filename>.e00* export files using the ARC/INFO import/export commands failed to correct the problem — whatever datastructure anomolies which developed in the original coverage were carried through the export files and continued to cause program crashes. The work-around which was successful involved using the application to generate EMME/2 batchfiles, and then rebuild a coverage from them.

Installation

The ***snapit.aml*** and ***movefwy.aml*** applications have been stored ***in /magtpo/net.***

2.7 Node Renumbering

As has been mentioned, the existing node numbering scheme displayed in MAG networks had evolved over the years so that (1) the actual values for node identifiers were more-or-less random with respect to geographic location and (2) node identifiers were not consistent across the networks. In addition to the geographic coordinate corrections desired by MAG, a more organized approach to the node (and TAZ) numbering scheme was desired. The three characteristics desired were as follows:

- Nodes appearing in the network were to be renumbered according to geographic location. The simplest scheme to implement involved numbering nodes as they appeared horizontally across the network.
- Node numbering should be consistent across all networks. That is, nodes which continue to appear in all networks should have the same identifier. Aside from convenience, EMME/2 software provides features to compare the assignment results of different networks — it depends on node numbers for the comparison (that is, the inode, jnode pair is the key to the link record).
- A node numbering scheme according to the rules set forth in Figure 2-7 was desired.

After initially renumbering the base year network according to this scheme, it was discovered that some of MAG's travel forecasting programs depend on the existing TAZ numbering scheme. Therefore, TAZ numbers were restored to their original values.

Figure 2-7
Proposed Node Renumbering Scheme

External Zones	1..16	Number clockwise starting at I-10/Buckeye
Geographic Locations	31..78	Add 30 to each of the existing geographic location identification numbers
TAZs	101..2000	Zone numbers should be grouped by geographic location and sorted by coordinate within each geographic location
Focused Zones	2001..3000	Reserve for future use
Network Nodes	3001..	Regular Network Nodes

The first objective, renumbering nodes in geographic sequence, was easy to achieve. Nodes were merely sorted based on their *x,y* location (actually, sorted on the y coordinate in descending value and the x coordinate in ascending value). This has the effect of numbering nodes horizontally across the region from the northwest.

However, the condition of the MAG networks made meeting the consistency objective difficult to achieve. As was indicated earlier, the only way to determine if two nodes appearing in two networks were in fact, the same, is either by the current node number or by its position. Unfortunately, neither could be assumed true in all cases. An attempt was made to retain node numbers as they migrate across the networks, as follows:

- The 1993 network was renumbered first, and progressed to future networks in order (the 1995 was second, the 2005 third, etc...).
- Nodes appearing in the 1993 network were sorted by their *x,y* coordinate and then renumbered in consecutive order (according to the record number). Then they were placed back into their original position in the file.
- Nodes in the 1995 network were then matched against their counterparts in the 1993 network, by position. This was accomplished by the ARC/ INFO *conflation feature* which permits points in one coverage to be matched against points in another coverage. A *low conflation tolerance* (100 feet) was used to prevent nodes appearing in two networks to be matched, when in fact they were not the same.
- The *original* networks were used as a basis for the renumbering, not the geographically adjusted networks. The original networks offered the best chance for nodes to be located in the same position, when in fact they represented the same intersection.
- In cases where matches were made, new node numbers were migrated from the previous network. In cases where new nodes appeared in the target network, new identifiers were assigned.
- An editing program was created (an ARC/INFO AML application ***renumber.aml***) which permitted the user to inspect nodes which had no counterparts in a previous network. The application let the user inspect *misses* to determine if the nodes in the target network were genuinely new, or in fact their location was beyond the conflation tolerance. One feature of the program was to let the user direct matches on an individual basis. The program would also allow the user to inspect nodes which matched based on position, when in fact the original identifiers were different (this would be a spurious hit). All of the networks were then personally inspected to attempt to clean up anomolies.
- Genuinely new nodes appearing in future networks were then renumbered according to the same x,y convention.

The results of this process were good, although not perfect. Lookup tables translating old node numbers to new node numbers were created. Figure 2-8 provides a summary report of the node number identifiers in each network.

Figure 2-8
Node Number Ranges after Renumbering

Network	Centroids	Node Number Range
net1993	1..1288	3001..9146
net1995	1..1288	3001..9750
net2005	1..1288	3001..10744
net2015	1..1288	3001..11797

The coverages generated from the renumbering effort are shown in Figure 2-9.

Figure 2-9
Network Coverages Created after Node Renumbering

Subject	Original Cover		Renumbered Cover
Rubber Sheeting	net93_fwy	»	net93_fin
	net95_fwy	»	net95_fin
	net05_fwy	»	net05_fin
	net15_fwy	»	net15_fin

Installation

The ***renumber.aml*** has been placed in the ***/magtpo/net*** directory. Equivalency tables associating old and new node identifiers for the networks were uploaded to the ***/consult*** directory.

Future Improvements

In retrospect, a different algorithm should have been used to generate node numbers, as follows:

1. First, the x,y coordinate values (or at least the y coordinate value) would have been rounded to the nearest 100-500 feet before sorting them. The problem with the technique that was adopted is that node number sequences along roadways in the network are *very sensitive* to the y-coordinate value, and therefore tend to skip around laterally across the map. Rounding coordinates to the nearest 100-500 feet would have the effect of numbering nodes laterally according to general geographic bands across the map.

2. Second, instead of building a cascading set of equivalencies to relate nodes in the 1993 network to the 1995 network, say, and then from the 1995 network to the 2005 network, it would be more efficient to dump all nodes from all networks into a single file, and then analyze them that way.

2.8 Metered Ramps

Initially, all of the editing of freeway ramps was done to match, to the degree possible, the underlying shape and configuration of the interchange ramp in the background network (without coding any new nodes). In many instances, especially for loop ramps, existing nodes were already coded in the network and therefore could be used as a shape point. Later, however, MAG indicated a desire to code *metered ramps* with 0-length. Metered ramps are *symbolic* and do not really relate to any real underlying physical feature.

As has been mentioned earlier, ARC/INFO representation prohibits arcs with 0-length and will automatically delete them from the coverage when encountered. Also, as a practical matter, 0-length links necessarily imply that both the inode and the jnode will be superimposed, presenting problems in finding and selecting them in either the ARC/INFO or EMME/2 environments. Consequently, metered ramps were coded to a minimal length calculated so that symbols (markers) for the inode and jnode would appear next to each other onscreen when the user zooms in on an interchange area. It was computed that an appropriate length to satisfy this objective to be 0.01 miles (about 50 feet).

However, revisions to metered ramps required us to modify network topology.

In order to do this expeditiously, a C-program (***meter.c***) was developed that reads and writes EMME/2 *batchin* files. The program finds links coded to be metered ramps (volume-delay functions which end in "8"), and moves the inode coordinate to be immediately adjacent to the jnode. The relocation of the inode must be done with some care, as the appropriate position may be above, below, to the right, or to the left of the jnode, depending on the direction of the ramp link. The program takes this into consideration and computes a new coordinate for the jnode. It also gives the resulting link the same overall compass bearing that it originally had. The results are then written to a new EMME/2 batch file.

The program is easy to run: it merely asks the user for the name of the batchin file to be modified and the output file desired. The program itself has been uploaded and stored in the ***/magtpo/net*** directory.

Files modified as a result of adjusting freeway ramp meter links are shown in Figure 2-10.

Figure 2-10
Network Coverages Created after Metered Ramp Adjustments

Subject	Original Cover		Adjusted Cover
Rubber Sheeting	net93_fin	»	net93_met
	net95_fin	»	net95_met
	net05_fin	»	net05_met
	net15_fin	»	net15_met

2.9 Turn Penalties and Transit Lines

Two other files associated with EMME/2 forecasting networks are keyed on node numbers, and therefore have to be revised. These are *turn penalties* and *transit line descriptions.* These files also required modification.

Turn Penalties: Turn penalty files in EMME/2 have a simple record structure: each record merely describes a three node sequence for each turn, along with the penalty associated with it. Since node number lookup tables were built in an INFO environment, a simple set of relates made it possible to easily modify turn penalty records.

Each turn penalty file was imported into INFO. New node numbers describing the turn were then computed from the lookup table. The resulting turn penalty table was then dumped to a new EMME/2 file. All of these operations were performed manually interactively at the terminal.

Note that turn penalty files were available for the 2005 and the 2015 networks *only.*

Transit Lines: Transit line descriptions could not be modified so easily, owing to the somewhat complicated record structure that EMME/2 uses. In fact, transit line descriptions do not really follow any type of fixed record format — instead transit lines are described in a kind of free format language of its own that must be *interpreted.* An example of a transit line description file is shown in Figure 2-11. Consequently, a special C-program (***transit.c***) was written to automate the process.

The key to recognizing node numbers in a transit line description file appears to be that any field (characters following a space) that *starts with a digit is a node number*; all other fields start with other non-digit alphabetic characters (this probably is how EMME/2 itself recognizes node paths). This is what ***transit.c*** does, any node that it encounters is replaced by its lookup table equivalent — everything else appearing in the record is left unchanged.

Figure 2-11

Example of EMME/2 Transit Line Descriptions

```
t lines init
a yellow b 1  30.00  12.00   'GRAND/ASU-YELLOW      '  4  201   1
 ttf=1 path=yes
 7200  dwt=+0.0  7201  dwt=#0.0  7204  dwt=#0.0  7212  dwt=+0.0  7093  dwt=#0.0
 7029  dwt=#0.0  6945  dwt=#0.0  6815  dwt=#0.0  6760  dwt=+0.0  6710  dwt=#0.0
 6593  dwt=+0.0  6589  dwt=#0.0  9223  dwt=#0.0  6552  dwt=+0.0  6546  dwt=#0.0
 6550  dwt=#0.0  6549  dwt=+0.0  6545  dwt=#0.0  6543  dwt=+0.0  6542  dwt=#0.0
 6540  dwt=+0.0  6537  dwt=#0.0  6515  dwt=#0.0  6531  dwt=+0.0  6521  dwt=#0.0
 6517  dwt=#0.0  6519  dwt=#0.0  6518  dwt=#0.0  6520  dwt=+0.0  6530  dwt=#0.0
 6533  dwt=+0.0  6532  dwt=#0.0  6529  dwt=+0.0  6528  dwt=#0.0  6525  dwt=+0.0
 6523  dwt=+0.0  6524  dwt=#0.0  6527  dwt=+0.0  6526  dwt=#0.0  6522  dwt=+0.0
 6516  dwt=#0.0  6512  dwt=+0.0  6500  dwt=#0.0  6501  dwt=#0.0  6503  dwt=+0.0
 6469  dwt=#0.0  6411  dwt=+0.0  6255  dwt=#0.0  6121  dwt=#0.0  6046  dwt=#0.0
 5833  dwt=+0.0  5757  dwt=+0.0  5674  dwt=+0.0  5543  dwt=+0.0  5393  dwt=+0.0
 5231  dwt=+0.0  5172  dwt=+0.0  5087  dwt=#0.0  4997  dwt=+0.0  4948  dwt=#0.0
 4897  dwt=#0.0  4803  dwt=+0.0  4711  dwt=+0.0  4618  dwt=+0.0  4574  dwt=+0.0
 4505  dwt=+0.0  4473  dwt=+0.0  4412  dwt=+0.0  4375  dwt=+0.0  4260  dwt=+0.0
 4209  dwt=+0.0  4153  dwt=+0.0  4210  dwt=+0.0
 lay=0.00
 4153  dwt=+0.0  4209  dwt=+0.0  4260  dwt=+0.0  4375  dwt=+0.0  4412  dwt=+0.0
 4473  dwt=+0.0  4505  dwt=+0.0  4574  dwt=+0.0  4618  dwt=+0.0  4711  dwt=+0.0
 4803  dwt=+0.0  4897  dwt=#0.0  4948  dwt=#0.0  4997  dwt=+0.0  5087  dwt=#0.0
 5172  dwt=+0.0  5231  dwt=+0.0  5393  dwt=+0.0  5543  dwt=+0.0  5674  dwt=+0.0
 5757  dwt=+0.0  5833  dwt=+0.0  6046  dwt=#0.0  6121  dwt=#0.0  6255  dwt=#0.0
 6411  dwt=+0.0  6469  dwt=#0.0  6503  dwt=+0.0  6535  dwt=+0.0  6536  dwt=#0.0
 6538  dwt=#0.0  6541  dwt=+0.0  6569  dwt=+0.0  6572  dwt=#0.0  6571  dwt=+0.0
 6564  dwt=#0.0  6570  dwt=#0.0  6568  dwt=+0.0  6573  dwt=#0.0  6576  dwt=#0.0
 6562  dwt=+0.0  6565  dwt=#0.0  6563  dwt=+0.0  6566  dwt=#0.0  6559  dwt=#0.0
 6560  dwt=#0.0  6561  dwt=#0.0  6567  dwt=+0.0  6515  dwt=#0.0  6537  dwt=#0.0
 6540  dwt=+0.0  6542  dwt=#0.0  6543  dwt=+0.0  6545  dwt=#0.0  6549  dwt=+0.0
 6550  dwt=#0.0  6546  dwt=#0.0  6552  dwt=+0.0  9223  dwt=#0.0  6589  dwt=#0.0
 6593  dwt=+0.0  6710  dwt=#0.0  6760  dwt=+0.0  6815  dwt=#0.0  7029  dwt=#0.0
 7093  dwt=#0.0  7212  dwt=+0.0  7204  dwt=#0.0  7201  dwt=#0.0  7200  dwt=+0.0
 lay=0.0
```

The program is easy to operate: it merely asks the user for (1) the name of the transit line file to modified, (2) the node number lookup table to be used, and (3) the output file that is desired.

Note that no transit line file for the 2015 network was available. Figure 2-12 summarizes the files that were created as a result of this process.

Figure 2-12

Renumbered Turn Penalty and Transit Line Files

Subject	Original File		Translated File
Turn Penalties	proh05f1.txt	»	turns05.new
	proh15f1.txt	»	turns15.new
Transit Lines	rteof93.txt	»	rteof93.new
	rtepk93.txt	»	rtepk93.new
	rteof95.txt	»	rteof95.new
	rtepk95.txt	»	rtepk95.new
	rteof05.txt	»	rteof05.new
	rtepk05.txt	»	rtepk05.new

2.10 Results

The geocorrection process conducted in this task has given MAG four updated travel forecasting networks for the years 1993, 1995, 2005, and 2015. These networks now conform to the Arizona State Plane Coordinate system and can be overlaid on other geographic coverages in MAG's growing inventory. Figures 2-13 and 2-14 illustrate examples of MAG travel forecasting networks after they have been geocorrected. In addition to these results:

- Networks are substantially cleaner than they originally were. Node identifiers have been reassigned and are consistent across the networks, which will make EMME/2 network comparisons easier to perform.

- Link lengths associated with the networks are now more accurate, so VMT estimates will also be more accurate.

Moreover, MAG now also have a capability to move travel forecasting datasets between ARC/INFO and EMME/2, opening the door to heavier reliance on GIS to support travel forecasting and analysis work in the future.

Figure 2-13
Example of Geocorrected MAG Network

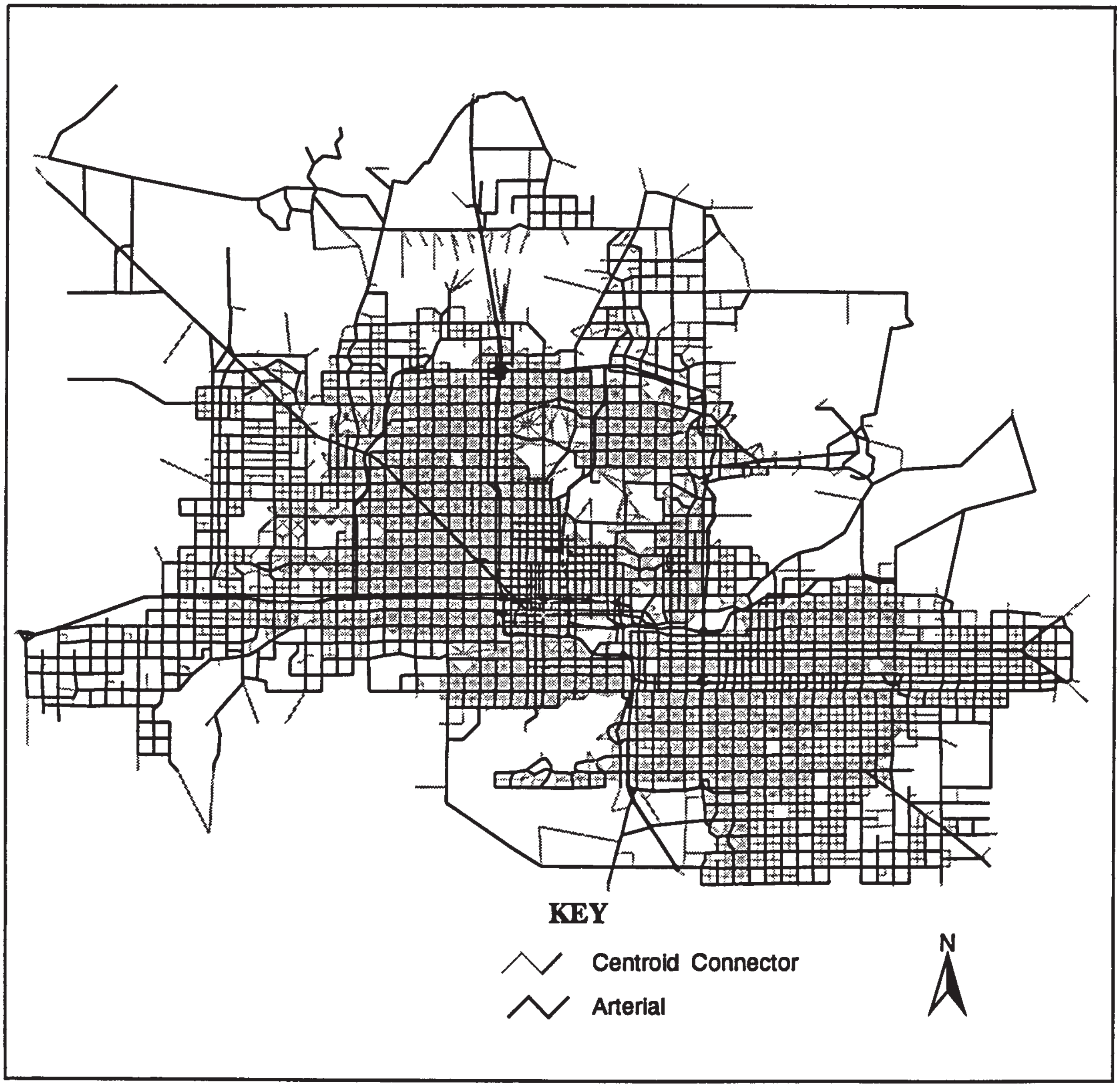
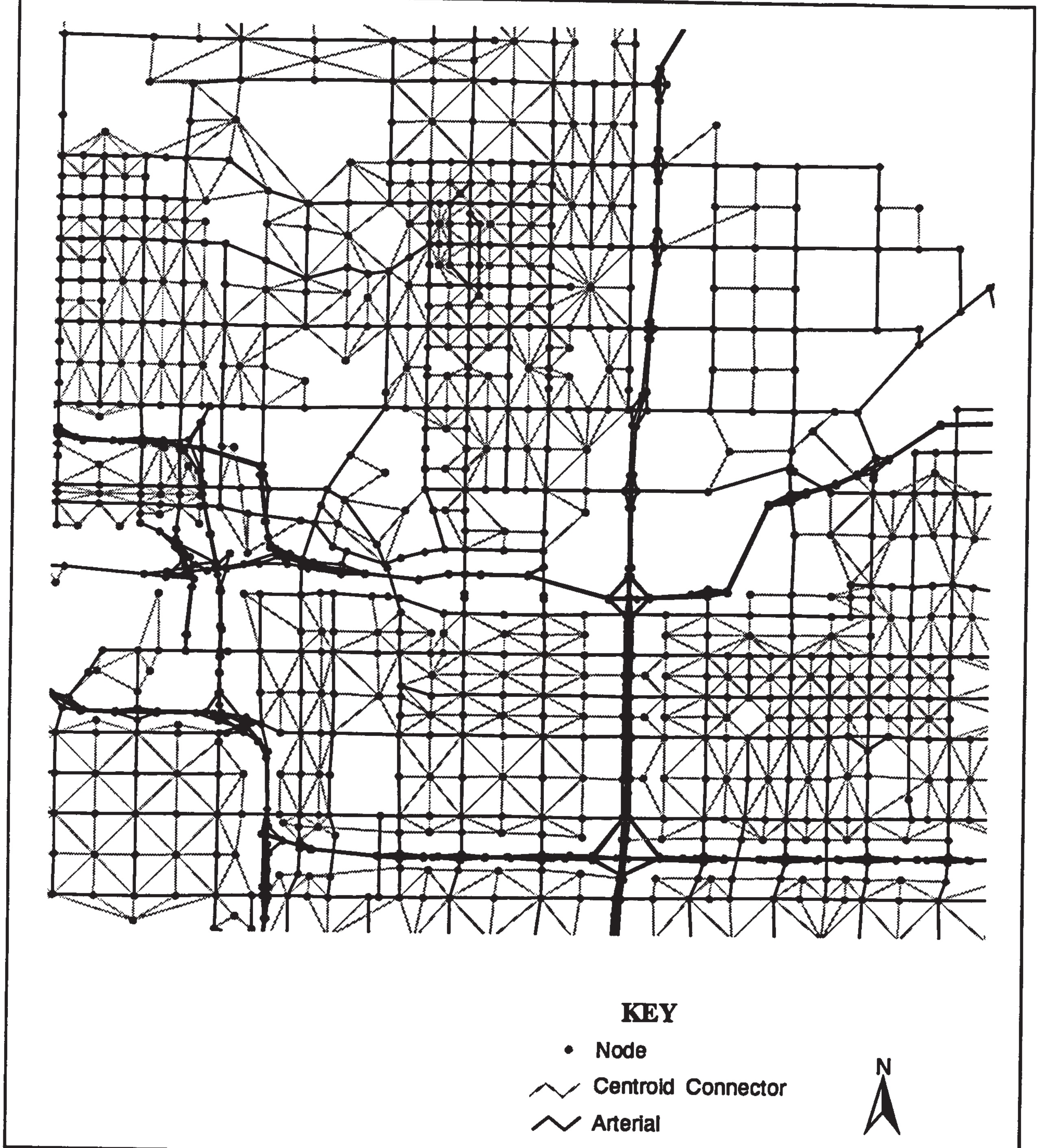


Figure 2-14
Example of Geocorrected MAG Network



3
Review and Update of MAG Geography

MAG has adopted several basic geographies on which much demographic, travel forecasting, and air quality analysis work is conducted. These structures are:

- The traffic analysis zone (TAZ) structure on which the EMME/2 travel forecasting model operates. Currently, there are 1,272 TAZs[1] in the modeling area. Figure 3-1 illustrates the TAZ structure now in use.

- The regional analysis zone (RAZ) structure is the geography on which the land use forecasting model DRAM/EMPAL operates. Currently, there are 141 RAZs in the region. Purely on the basis of the overall regional average, each RAZ encompasses roughly 9 TAZs. Figure 3-2 illustrates the RAZ structure now in use.

- The Metropolitan Planning Area (MPA) zone system is a summary level geography developed by MAG which loosely conforms to municipal boundaries. There are 35 MPAs covering Maricopa County, so on the average each encompasses roughly 4 RAZs and 36 TAZs. Figure 3-3 illustrates the MPA structure covering the County.

The TAZ structure is upwardly compatible with the RAZ structure (meaning that TAZs aggregate to RAZs) except for the fact that the RAZ structure is broader in scope geographically. While the RAZ structure covers all of Maricopa County, the modeling area does not. Equivalency tables are available to relate the RAZ structure to its TAZ members. Similarly, the RAZ structure is upwardly compatible with the MPA structure, and equivalency tables are available to list RAZ membership in each MPA.

---

[1] The travel model also includes 16 external stations which are treated as TAZs, so there are 1,288 zones in the current implementation. Also, coincidentally, these are 16 TAZs defined *outside* of the modeling area which are not used in the forecasting chain, although they are in Maricopa County.

Figure 3-1
Traffic Analysis Zone (TAZ) Structure in the MAG Region

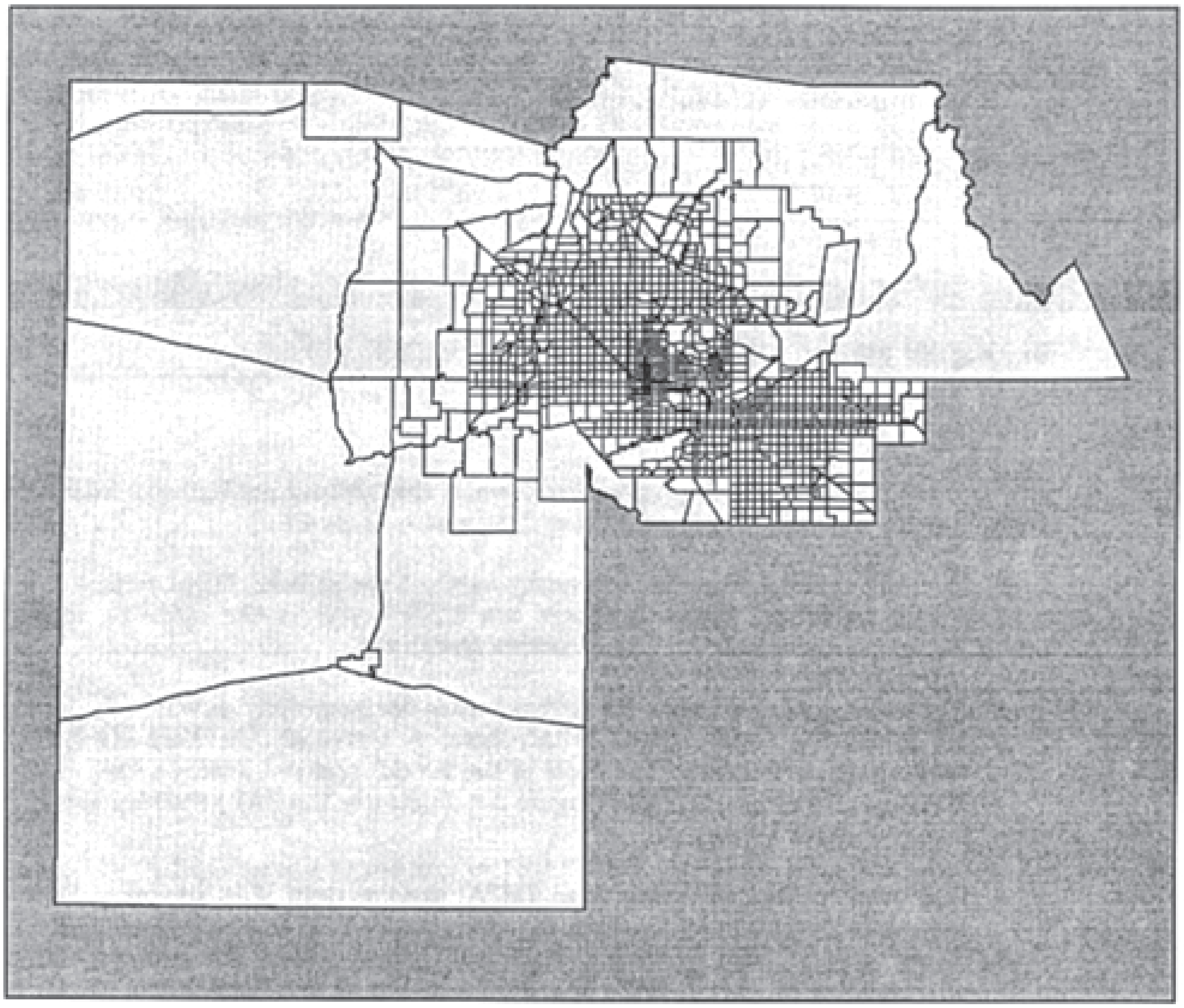


Figure 3-2
Regional Analysis Zone (RAZ) Structure in the MAG Region

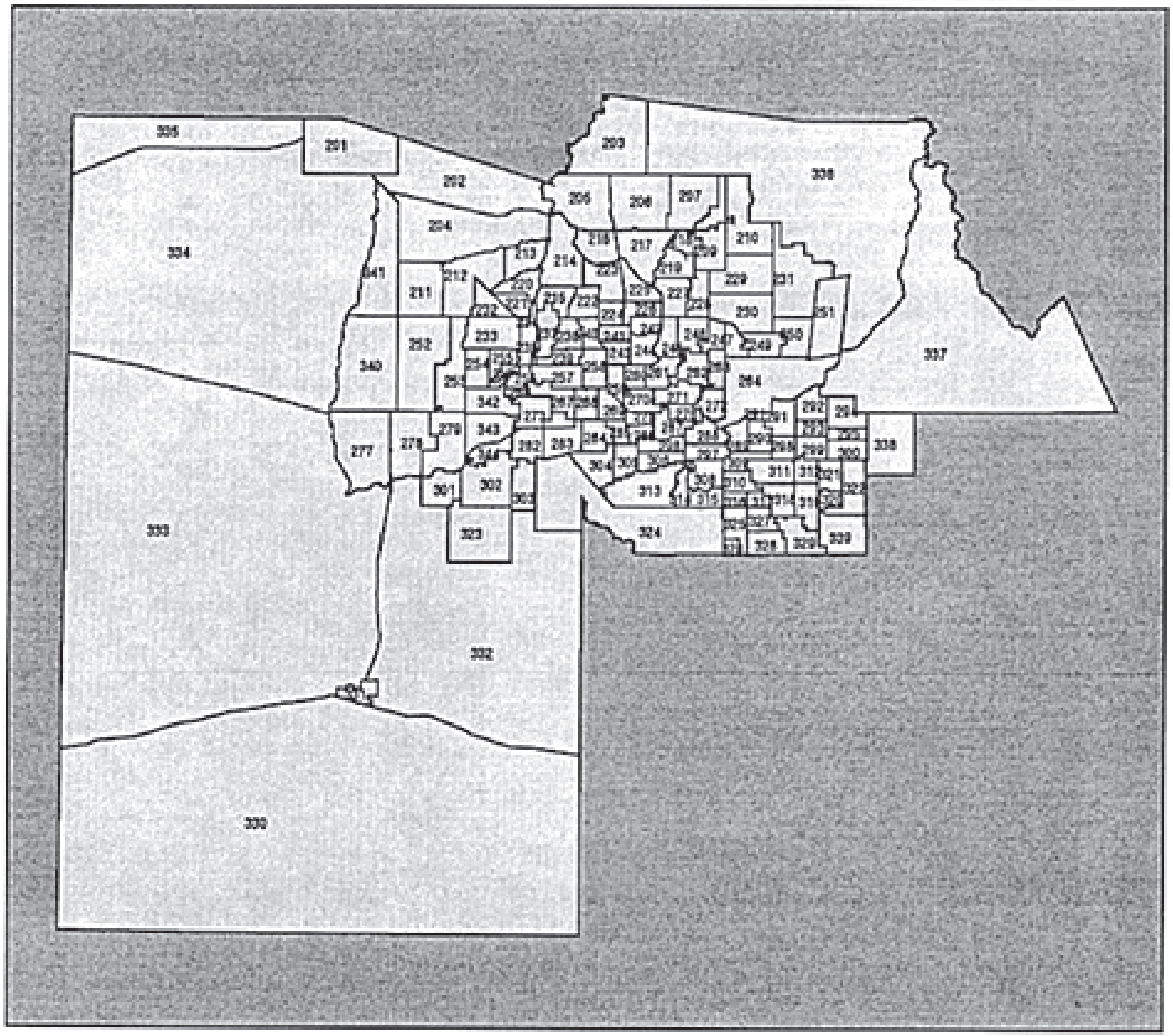
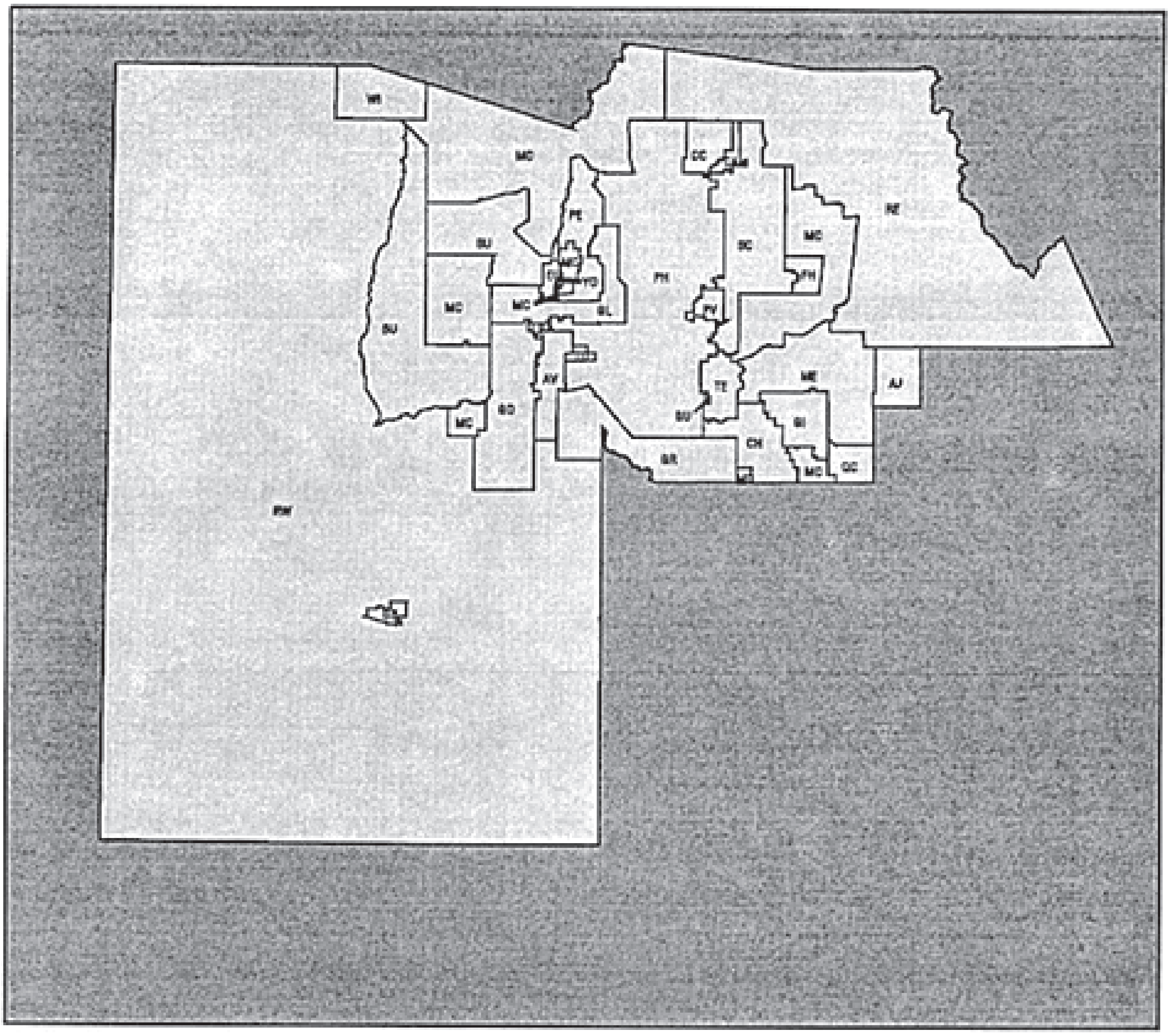


Figure 3-3

Metropolitan Planning Area (MPA) Structure in the MAG Region



Routine analysis operations at MAG, however, involve other geographies as well:

- The Census is a major source of demographic information, and therefore various census geographies (census block, block group, and tract) are needed to supply estimates for various socio-economic variables to MAG's system of databases.
- There are a wide variety of political boundaries in use in the region, including municipal boundaries, water districts, etc. MAG is frequently called upon to report demographic and model information for these geographies.

3.1 Task Objectives

A continuing problem that MAG analysts have to contend with is the bewildering variety of geographies on which source data is obtained and analysis data is desired. While MAG does an excellent job designing its own geographies to be consistent and compatible, there is no escaping the fact that substantial incompatibilities will inevitably occur. The Census Bureau, for example, designs census geographies according to its own criteria which apply nationally and have nothing to do with travel forecasting and other local needs. And political boundaries, of course, are established for purposes completely divorced from statistical analysis.

Geographies of different levels of detail present analysts with a continuing need to aggregate and disaggregate data — the aggregation problem is quite trivial when geographies are upwardly compatible, but the disaggregation problem can be difficult. When geographies are *not* compatible at all, then both aggregation and disaggregation can present considerable problems.

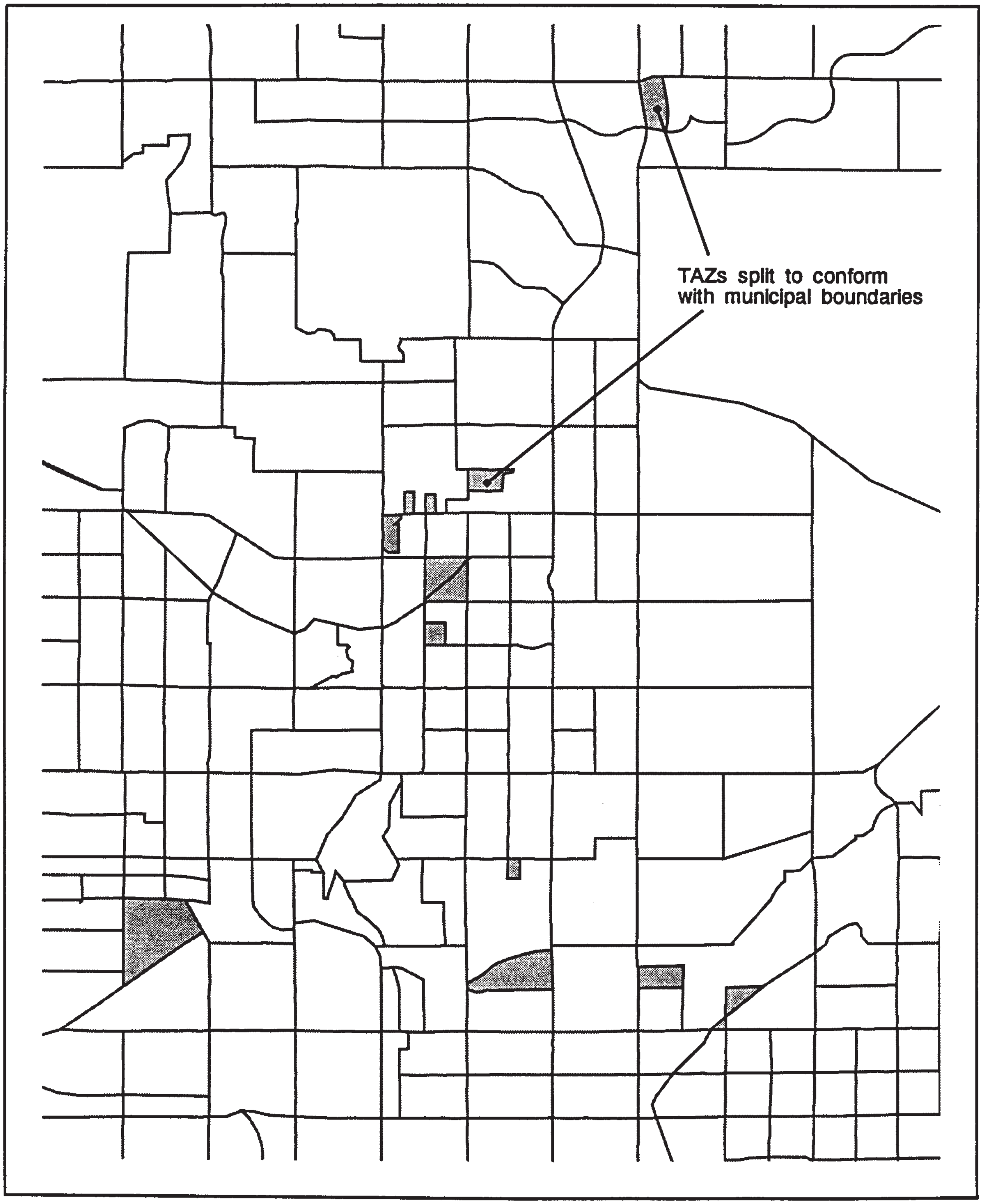
An attempt was made in 1993 [2] to deal with this problem through a proposed revision to the TAZ structure (and by implication, the RAZ and MPA structures as well). The basic thrust of this revision involved splitting TAZs so that the resulting boundaries could be aggregated to census tracts and municipal boundaries. In retrospect, however, this approach was not entirely satisfactory, as travel modeling really requires certain conventions to be observed. **Note that this TAZ structure was never put to use.** MAG continues to use the 1,272 TAZ geography that was developed a number of years ago.

In order for the transportation models to work properly, TAZ shapes should be somewhat regular and should conform with the underlying transportation network. The effect of splitting TAZs in order to force congruence with political boundaries results in the generation of numerous odd-shaped zones not necessarily bounded by streets. Figure 3-4 illustrates an example of this in a small section of Maricopa County. Another approach to this problem was explored in this task.

---

[2] See the Socioeconomic Models Enhancement: Final Report by the Economic Strategies Group; May 1994. This TAZ system had 1,522 zones, and has been uploaded to MAG as /magtpo/taz/taz_esg.

Figure 3-4
An Example of Unsuitable TAZs



The work in this task address two principal objectives:

- Development of a *new* geography that would facilitate MAG's requirements for dealing with numerous *other* incompatible geographies. The idea would be that this new geography, called a **Minimal Analysis Zone (MAZ)**[3] structure would be at a very fine level of detail and could be upwardly aggregated to any other geography in use in the region (including TAZs). Basic land use forecasting data would be disaggregated to this level of geography with the use of a **SubArea Allocation Model (SAM)** later in the project and would therefore effectively replace the use of TAZs which currently serve this purpose.

- Reconsideration of the TAZ geography on which the travel forecasting process is based. The new TAZ structure would be designed only to serve the purposes of supporting travel forecasting.

3.2 Traffic Analysis Zone Structure

Once relieved of its responsibilities to provide the basis for statistical analysis on a variety of geographies, a new TAZ structure could be designed purely in response to travel forecasting needs. Some of the design criteria which apply to the TAZ structure include:

- *Regularly shaped TAZs:* In a travel model, activity (e.g., trip ends) in a TAZ are all represented to take place at a single location — the TAZ centroid. Access to the transportation network bounding the TAZ is facilitated through symbolic network links called *centroid connectors* or *loadlinks*. Oddly shaped TAZs make the identification of a true centroid (that is, the center of activity) difficult (if not impossible). They also make the coding of access connectors to the network difficult.

- *Uniform Sizes:* The convention in travel models is that TAZ sizes should conform loosely to the prevailing street network and land use densities in a region. This is partly because there are practical limits to how many TAZs can be used — MAG's current EMME/2 license, for example, does not permit more than 1,600 zones. Typically, zone sizes may run anywhere between 1-block to 1-quarter mile in densely developed areas such as downtowns. Elsewhere, TAZs are typically 1-mile square, conforming to the 1-mile spacing of street grids typically found in suburban areas. Zone sizes in outlying rural areas are larger and are generally derived from the geographic features and rural highways that are found there.

- *Uniformity with Network Detail:* Models work best when TAZ sizes are consistent with link spacing in the underlying travel network. That is, the model

[3] Land use forecasts at MAG are currently performed on a RAZ basis using the forecasting software DRAM/EMPAL. MAG disaggregates DRAM/EMPAL forecasts to TAZs, and so TAZs now serve the same purpose that MAZs are proposed to serve here.

works best when TAZs are *bounded* by network links (or by geographic features such as rivers, etc.). Model systems that involve a gross mismatch between network density and the granularity of the TAZ system do not load very well.

The 1993 TAZ proposal violated all of these considerations, because the TAZs were designed for the purposes of anchoring an ability to aggregate socio-economic data, not travel forecasting. **The current 1,272 zone TAZ geography currently used by MAG in EMME/2, on the other hand, *does* observe all of these criteria.**

Still, an updated TAZ geography is needed by MAG for future use. Among the objectives of this new geography were:

- The current TAZ structure does not extend to all of Maricopa County. In the future it will, so new TAZ zones would be needed on the periphery of the county.
- A desire to renumber TAZs (that is, a new set of TAZ identifiers) to conform with a new numbering scheme. The new scheme is shown in Figure 3-5, and allocates series of zone identifiers for other local governments to use (e.g., cities) when generating focused models based on the regional model at MAG (see Section 3, where this issue was also addressed).

Figure 3-5

Proposed Node Renumbering Scheme

External Zones	1..16	Number clockwise starting at I-10/Buckeye
Geographic Locations	31..78	Add 30 to each of the existing geographic location identification numbers
TAZs	101..2000	Zone numbers should be grouped by geographic location and sorted by coordinate within each geographic location
Focused Zones	2001..3000	Reserve for future use
Network Nodes	3001..	Regular Network Nodes

In response to these objectives, the following was developed:

- A new 1,760-zone TAZ system extended to cover all of Maricopa County. The design of this system observed the criteria outlined above. However, the addition of new TAZs outside of the existing modeling area exceeds MAG's current EMME/2 license for 1,600 TAZs and therefore cannot be applied until MAG upgrades the license (which they might consider doing sometime in the future). The coverage which represents this zone system has been uploaded to MAG and now resides on ***redmtn*** as ***/magtpo/taz/taz1760.***

- Therefore, another 1,558-zone TAZ system was developed which conforms to MAG's current EMME/2 license. This was created by providing less detail in many outlying areas which are currently undeveloped or of very low density. Roughly 42 zones are available within MAG's current license for focused studies[4] that might be undertaken in the future. The coverage which represents this zone system has been uploaded to MAG and now resides on ***redmtn*** as ***/magtpo/taz/taz1558.***

- A new numbering scheme for the *existing* zone system was developed which observed the objectives outlined by MAG. An equivalency table (and coverage) relating old TAZ identifiers to new TAZ identifiers has been uploaded to MAG and now resides on ***redmtn*** as ***/magtpo/taz/taz90_renum.***

Several workshop sessions were held to review the results of these efforts with MAG staff. Currently, MAG's model development efforts are heavily oriented to successful implementation of the *existing* regional forecasting model on EMME/2. As a practical matter, the introduction of these significant changes in the underlying zone system at this time would represent a substantial effort. Therefore, further work in this area was deferred to the future.

There will be several additional work items which will be required to make any of these recommended TAZ geographies operational in the model:

- *Polygon Construction:* The coverages that was developed are *arc* coverages, not *polygon* coverages. An editing assignment will be required to *build* the coverages as polygons and to assure that they observe polygon topology. It is not worth doing this until MAG is satisfied with TAZ structure.

- *Loadlinks:* Loadlinks in all of the networks which will operate under the new TAZ geography will have to be added or edited.

- *Trip Generation Datasets:* The forecast socio-economic datasets used by trip generation will have to be revised to observe the new TAZ geography.

- *Trip Generation Programs:* Some revisions to the suite of trip generation programs will be required to accommodate the new TAZ geography.

- *Externals:* Extension of the modeling area usually means that external cordon stations presently used will be relocated also, requiring the need to obtain traffic

---

[4] In a focused study, the modeler typically windows in on a specific area of interest and disaggregates TAZs there to obtain a fine level of detail, thereby creating more TAZs in the region. Sometimes, TAZs in distant places are aggregated to offset the total count.

counts at the new locations and revising the treatment of externals in the current model chain.

Examples of the final recommended TAZ structures are shown in Figure 3-6 and Figure 3-7.

3.3 The MAZ Structure

As mentioned, aggregation of data from a finely grained geography (*microscale*) to a more coarsely developed geography (*macroscale*) is a trivial problem when the geographies are upwardly compatible — a simple lookup table enables the analyst to accumulate subtotals for whatever demographic parameter is needed.[5] Clearly, if *all* MAG data were maintained at a microscopic scale that was upwardly compatible with all other possible zone geographies, then statistics for any geography could be easily generated. Since the source for the data of interest is DRAM/EMPAL, which operates on RAZs, DRAM/EMPAL forecasts would only have to be disaggregated to this microscopic scale *once* during the forecasting process. After that, MAG could generate statistics on whatever geographic scale is desired.

With the TAZ geography question settled, the other objective in this task was to develop a fundamental microscale geography that MAG could use as a source for socio-economic data. This geography is called the Minimal Analysis Zone system, or MAZ.

There are two approaches available in GIS to generate the MAZ structure:

- The MAZ could be considered to be the union of all geographies of interest.
- Some GISs (specifically ARC/INFO) provide another dramatically different possibility called *grid*. ARC/INFO grid provides a capability to convert *vector* representations of geographies (polygons) into a *raster* representation (geographic cells).

A discussion of these will be covered next.

The Polygon Union Approach

A zone system could be invented that consists of the *union* of MPAs, RAZs, TAZs, census blocks, municipal boundaries, water districts, school districts, and any other possible active geographic zone system considered important enough for demographic summaries to be required. In theory, GIS provides a capability to overlay *all* of these zone systems on top of each other in a series of layers and to generate a single polygon coverage consisting of the union of all of them. The resulting MAZ structure would be the target geography for the Subarea Allocation Model (SAM).

---

[5] It's trivial for absolute quantities anyway. It's not necessarily trivial for statistical parameters such as means, such as average income.

Figure 3-6
Restructured MAG TAZs

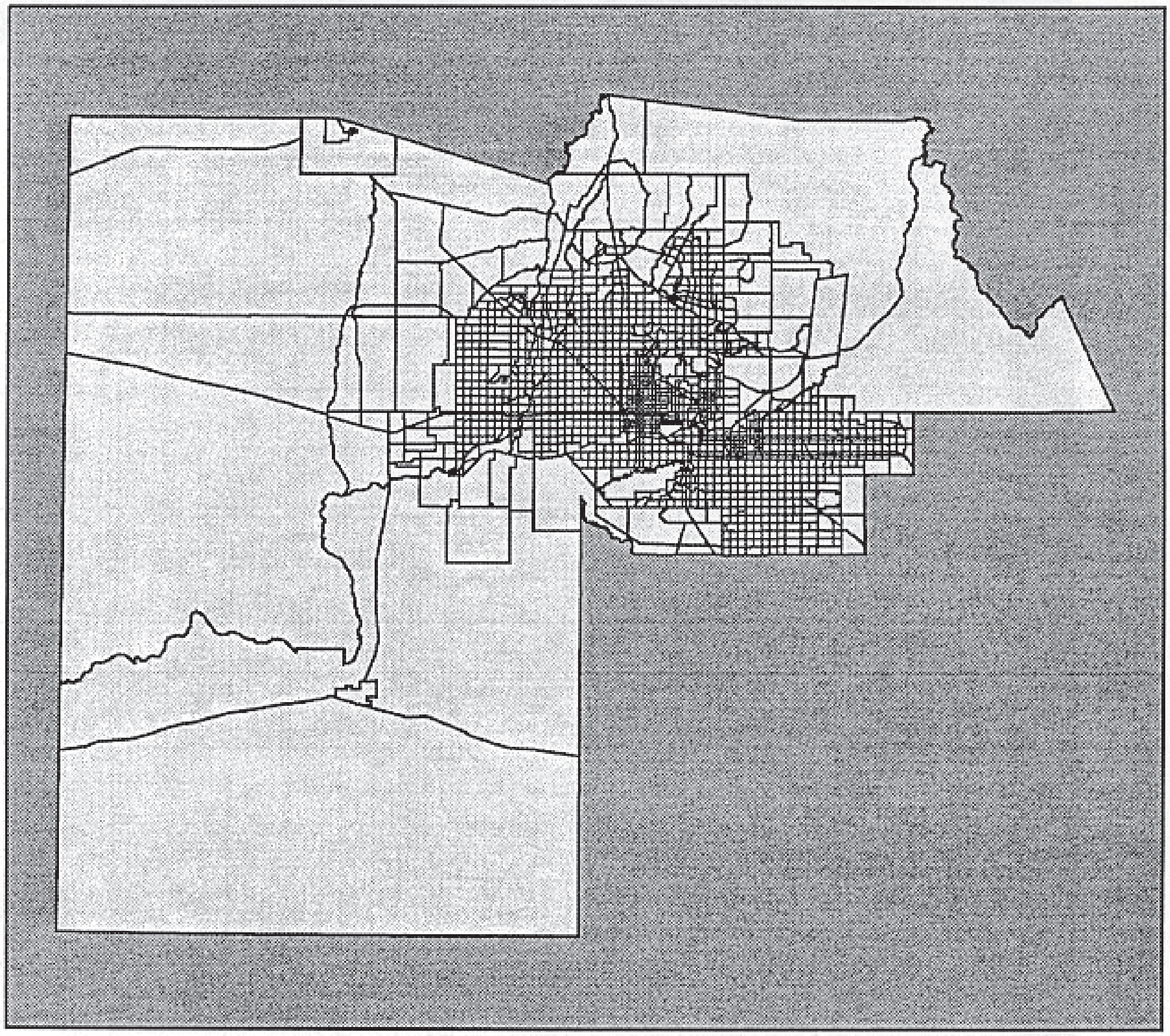


Figure 3-7
Blow-Up of Restructured MAG TAZs

Unfortunately, the brief trial investigation of this approach proved it to be impractical. Here's why:

- The union of many polygon coverages typically results in a polygon coverage full of *slivers*. The *sliver polygon problem* is a common and well known issue in GIS. A sliver is a spurious polygon that appears because of the geographic inaccuracies of the polygons which were overlaid to create it. (See Figure 3-8 for an example). For example, the overlay of a TAZ coverage that was digitized from street centerlines with a water district coverage that was digitized from parcel files would result in the creation of numerous polygons that might be 1 mile long and 33 feet wide — the difference between street centerlines and right-of-way lines. While the GIS would be correct in doing this, it certainly is not what was intended.

- Even when polygon coverages ostensibly are derived from the same geographic features (e.g., street centerlines), slivers will appear. This is because the coverages vary in accuracy. With all of MAG coverages projected to Arizona State Plane Coordinates (the units are feet). A digitizing or resolution error of 1-foot in one of the coverages will result in the generation a 1-foot wide sliver polygon when the covers are unioned. MAG has established a library of a wide variety of coverages from street centerlines, TIGER, from other local agencies, all of which were developed at different scales and resolutions. While the coordinates associated with them are *precise* (that is the coordinates are double-precision floating point binary numbers), they are certainly not of the same *accuracy*.

- The only way that slivers can be avoided is if the polygon coverages which contribute to the union were all built from the same source — for example, that a municipal boundary cover was built by extracting relevant arcs from the TAZ cover.

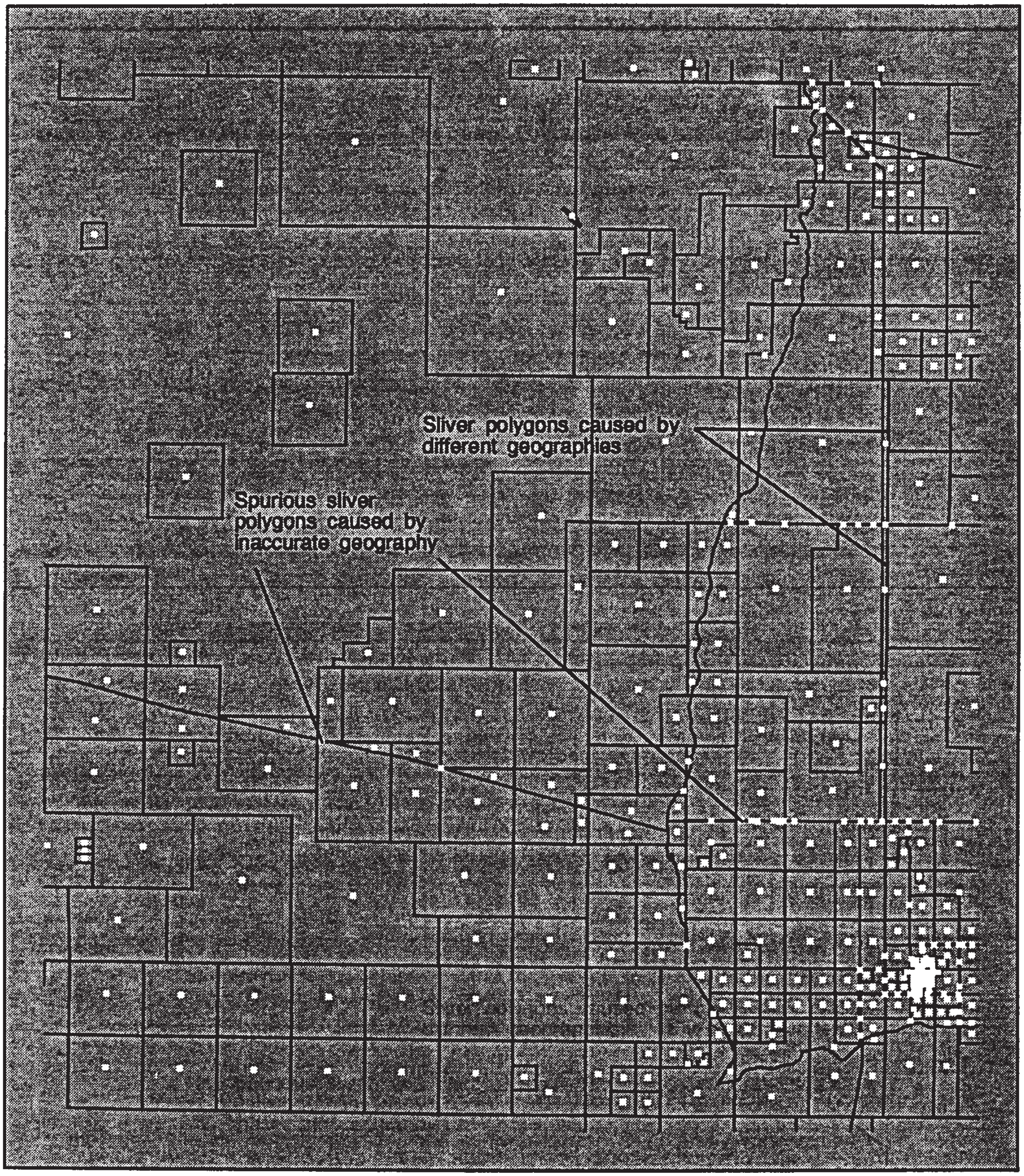
The sliver problem is quite serious. Not only does the resulting MAZ contain numerous polygons that, in fact, represent nothing, but ARC/INFO can also have significant difficulties dealing with them. During the trial, system crashes were generated when ARC/INFO was asked to perform even simple operations on the trial MAZ structure created from the union of others.

In order to use this union approach all MAG polygon coverages have to be created with the exact same set of arcs and coordinates. For example, the MAZ would have to be built from arcs that actually appear in all of the other coverages. Once established, all of the other coverages would have to be regenerated from the MAZ.

An approach based on polygon unions is impractical for other reasons as well:

- It is common in many regions that municipal boundaries, school districts, and water districts change over time. A MAZ structure built on unions of these boundaries will require high maintenance — every time there is any change in any of the contributing geographies, everything connected with the MAZ will

Figure 3-8
Example Sliver Polygons Appearing in a Union MAZ Dataset



have to be regenerated (including the disaggregated socio-economic data associated with it).

- Moreover, a MAZ built on polygon union assures MAG of the capability of generating data only to those geographies on which it is based. If MAG receives a request to generate statistics on a geography that was not anticipated at the outset, then a new MAZ structure will have to be created or an external approximation technique used to produce the requested data.

The Grid Approach

ARC/INFO provides another technique which is much more *robust*. This technique is known as *grid*.

Grid is a cell-based representation of geography.[6] Any region can be simply gridded into thousands of small square cells at any resolution that is desired. For example, all of Maricopa County can be divided into 220-foot square cells (a little over 1 acre in size), and these cells can be considered to comprise the MAZ structure for the region. ARC/INFO provides many features for transforming point, arc, and polygon features into grids and vice versa, as well as commands for manipulating, computing, and analyzing grid geographies.

A grid approach does offer several disadvantages:

- There is unquestionably a loss of precision with grids. Individual grid cells must belong to one polygon or another, regardless if they are bisected by the boundary. Quite apparently the loss of precision decreases with an increase in resolution.

- Just as the case with raster images, geographic grids imply much greater storage requirements and computational time because of the large amount of data that results. Still, ESRI has devoted much attention to this problem and various algorithms for reducing the size of datasets (such as run-length encoding) and increasing computing speed make grids a practical solution, even for as large a region as Maricopa County.

And, an approach to the MAZ predicated on grids offers many advantages:

- The MAZ structure can be generated on the fly, because it is based purely on the geographic extent of Maricopa County. It is not based on any other underlying polygon coverage.

- This means that TAZ covers, municipal boundaries, and water districts can change over time with no effect on the MAZ structure or MAG's ability to aggregate to the new geography.

---

[6] Clearly, the notion of a geographic grid comes from the notion of raster imagery. Like early graphics systems, traditional GIS representation of features is in terms of vectors. Today, many graphics systems (e.g., X-windows, image processing, etc.) find more power with raster representation.

- Also, this means that MAG can generate statistics from the MAZ for geographies that were not anticipated at all at the outset.

- MAG does not have to maintain equivalency tables with an approach based on grids. Aggregation to higher level geographies is done based on cell position in the grid.

- Some forms of geographic analysis which are difficult to perform on polygon coverages with traditional analysis techniques are easy to implement on grids (e.g., measuring accessibility to neighboring land uses represented by grids).

To test the concept, a number of grid structures were implemented for Maricopa County and sample MAG datasets. This subject is discussed next.

3.4 MAZ Implementation

A number of benchmark trials were run to test the application of ARC/INFO grid data structures as they would apply to the MAZ concept, as follows:

- A number of grid structures were generated representing TAZ datasets for various grid cell sizes, ranging in resolution from 1 mile all the way to a microscopic scale of 110-feet (approximately a quarter-acre grid cell).[7] The processing time was measured and storage space to represent all of Maricopa County this way. In general, higher resolution grids would be preferred to improve accuracy, however processing time and data storage requirements will increase.

- In addition, an AML-application was generated which gives the user the capability to use the grid concept for aggregation and disaggregation. The test case involved 1995 socio-economic datasets for TAZs and RAZs. The benchmarks involved processing time and accuracy as a function of the grid-cell size selected. (The AML-application will become part of the Subarea Allocation Model covered in Chapter 8.)

These tests provide some insight into the power associated with ARC/INFO grid and the ability of grid to serve as the MAG MAZ structure.

---

[7] MAG originally built the land use and general plan coverages from grid representations using 440-foot size grid cells (approximately 4.4 acres).

Grid RAZ Benchmarks

MAG's 1995 socio-economic dataset reported for TAZs was gridded for a single variable, household population for a series of grid sizes ranging between 1 mile and 110-feet.[8] These tests were run on a 90 *mhz* Sun SPARC-Station 10 running Solaris 2.3 and ARC/INFO Version 7.01. Since the file structure existed on a remote file system, processing times reported here also reflect network bandwidth. (MAG's own results on the Hewlett-Packard workstations may be different, but the comparison is intended to address order-of-magnitude expectations and relative performance based on grid-cell size).

The results are shown graphically in Figure 3-9, and can be summarized as follows:

- Disk storage requirements and processing time vary with respect to the square of the size of the grid (in other words, grid is a *quadratic* or *parabolic* order technique).

- The *knee* of the curves shown in Figure 3-9 all show break points at grid sizes of about 220-feet (approximately 1-acre grid cells). That is, the grid resolution can be increased from 1-square mile grid cells all the way down to 1-acre grid cells without paying much of a penalty in terms of storage or processing time. From that point on, higher resolution carries increasingly heavier costs in terms of storage and processing time.

- Even so, both storage and time requirements associated with generating a grid dataset from a TAZ coverage is remarkably acceptable. Even a very microscopic-sized cell of 110-feet on a side (Over 33 million quarter-acre grid cells for the entire County) was generated in just under 3-minutes and consumed less than 1 megabyte of storage.

More details about how ARC/INFO Grid works and how files are organized is available from ESRI documentation. An important consideration is that for efficiency purposes, ARC/INFO does *not* store values for every cell in the grid — it *only* stores unique values.

Therefore, a 110-foot grid that holds only 1,272 unique values (say, population for each TAZ) will require only 1,272 records — not 33 million records associated with each individual grid cell.

Aggregation and Disaggregation

In addition, grid procedures were tested for use in aggregation (and disaggregation) by developing an application that takes as input:

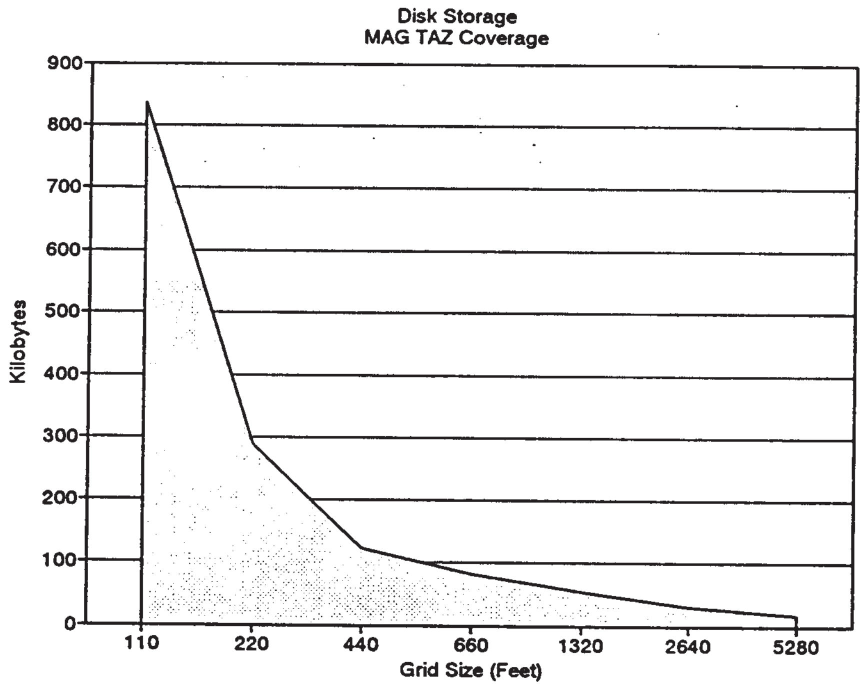
- A polygon coverage with information to be aggregated (e.g., TAZs)

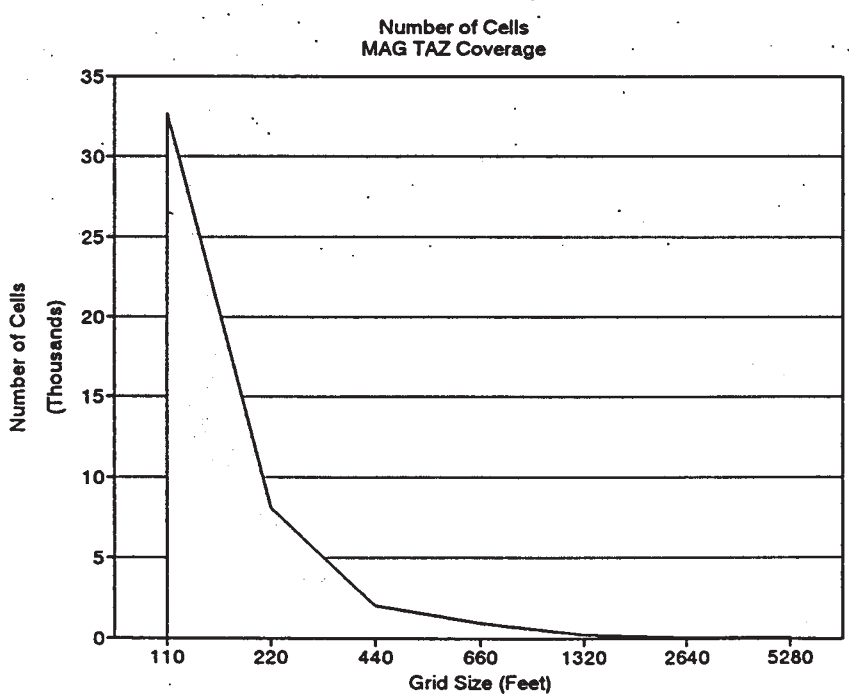
---

[8] It would seem that grid cell sizes should be multipliers of 5,280 feet so that they conform to the geography of section lines and street grids and other features which follow regular spacing intervals.

Figure 3-9

ARC/INFO Grid Performance Associated with Cell-Size





- A target polygon coverage to which information was to be aggregated (e.g., RAZs)

The application aggregates the information. The results of time trials in this *proof-of-concept* benchmark are shown in Figure 3-10. As noted there, timings ran on the order of 12 minutes for even the most microscopic scale geographic of 110-foot grid cells. These times *include* the time required to generate the grids themselves for both covers — the actual aggregation summations are relatively fast.

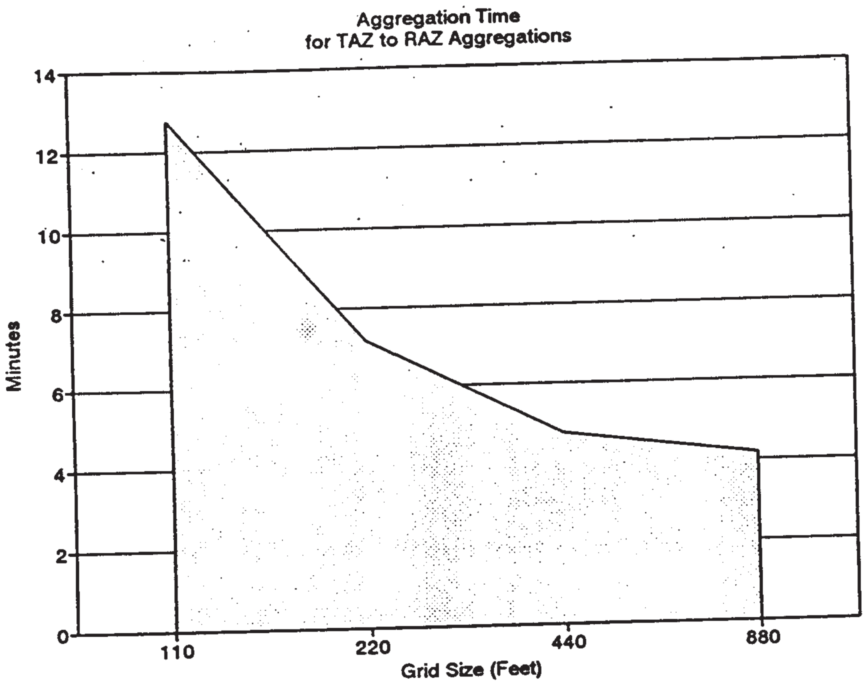
Also, note that the correct answer from the aggregation is known, because MAG maintains lookup tables which can be used to generate aggregate RAZ data from TAZs. In the benchmark test it was found that the grid result was also *exactly* correct. This is because the RAZ geography was built from the TAZ geography and they are therefore exactly compatible. Exact results should not be expected when aggregating TAZ coverages to, say, municipal boundaries. Then, the resolution of the grid size will play a role in the accuracy of results. (However, no prototypical data sets to use to test the relative error associated with grid resolution).

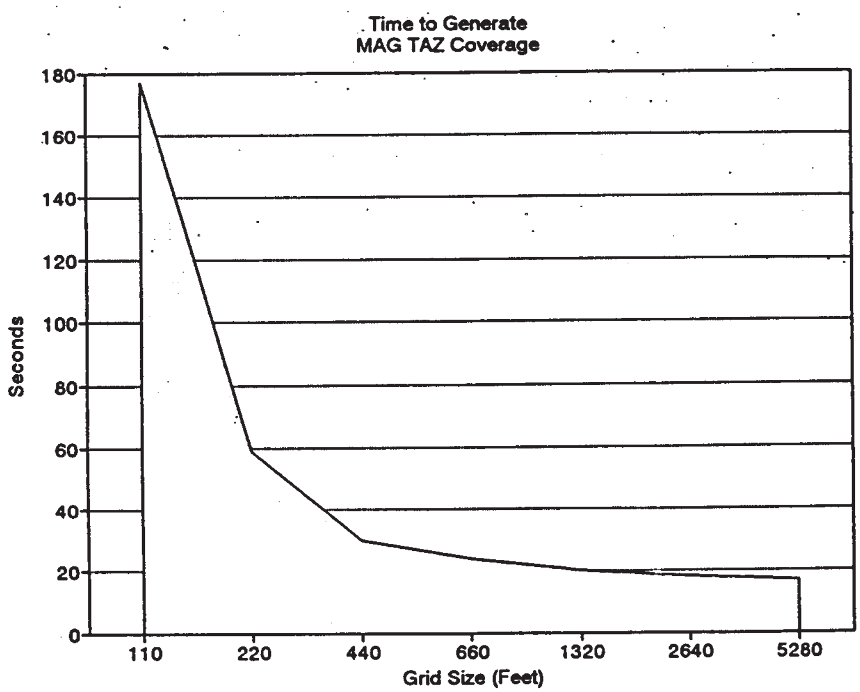
Conclusions

ARC/INFO grid gives every indication of providing an effective tool for managing socio-economic data for a variety of geographies — it was found to be efficient in terms of both processing time and storage requirements. Therefore it is going to rely on ARC/INFO grid as a basis for the MAZ structure. The actual implementation of the MAZ and the algorithms and applications associated with it will arise in Chapter 8. The MAZ dataset will be created then and the SubArea Allocation Model (SAM) will be predicated on it.

Figure 3-10

Times Associated with Aggregating TAZ Data to RAZs Using GRID





4
Review of MAG and ESG County-Level Population Projections

4.1 Executive Summary

Under Tasks 5 and 7 of the GIS Analysis, and Socioeconomic and Transportation Data Enhancement Project, the Applied Management and Planning Group (AMPG) undertook a review of the Maricopa Association of Governments Transportation and Planning Office's (MAGTPO) derivation of county population and household estimates for five-year intervals through year 2040, and the Economic Strategies Group (ESG) estimation of county-level special population. MAGTPO used the Arizona Department of Economic Security (DES) population projections as the basis for all of its estimates, which are in turn used as inputs to the DRAM/EMPAL land use forecasting models.

Minor unrealistic assumptions were identified in the estimation of employment totals and employment sector shares. However, MAG's estimates of special population and housing units deviated greatly from ESG's projections. Differences in definitions might have caused these discrepancies. Because of insufficient information, these differences could not be resolved.

Upon review of MAGTPO's procedures, AMPG recommends the following enhancements:

- Modification of wage and salary equivalence ratio to attain reasonable employment figures relative to DES year 2002 projections;
- Modification of the DES employment share matrix to provide stable estimates for the Mining sector;
- Addition of a procedure to match employment totals after DES sector employment shares are converted into MAG sector shares;

- Performance of sensitivity analysis to assess the effects of MAG and ESG estimates of housing units for special populations on the DRAM/EMPAL models.

Examination of ESG procedures resulted in the following recommendations:

- Overall, the base population used in most projections should be reviewed and updated using DES projections.

- AMPG recommends a further study on the occupancy rates for the group quarters population to estimate housing units occupied by the group quarters population.

- ESG estimation of the domestic transient population was sufficient. Projections on foreign visitors relied heavily on population estimates of selected countries, particularly Mexico, and did not consider other important economic factors. Alternative methods should be devised.

- ESG used an identical percent share of visitors from different states and Canada for estimating mobile home and recreational vehicle dwellers. Seasonal population in other housing was dependent on MAG's projection of total housing units, assuming a fixed percentage of total housing to arrive at available housing and assuming 2.0 individuals per unit. The last assumption might have created the unusually large seasonal population in other housing.

4.2 Introduction

The Applied Management and Planning Group (AMPG), under sub-contract to Barton-Aschman Associates, Inc., evaluated the methodologies used by the Maricopa Association of Governments (MAG) in deriving population estimates at five-year interval period from 1995 to 2040. Under Task 5 and portions of Task 7 of the GIS Analysis, and Socioeconomic and Transportation Data Enhancement Project, AMPG reviewed the existing procedures used by MAG Transportation and Planning Office (MAGTPO), and the projections and procedures suggested by Economic Strategies Group (ESG) in their final report, "Maricopa Association of Governments Socioeconomic Models Enhancements." MAGTPO furnished AMPG with Excel spreadsheets that automate MAGTPO's forecasts.

The following section briefly describes the procedures currently being used by MAGTPO to generate county population and household totals.

Procedures Deriving Control Totals

Under the State of Arizona Executive Order 95-2, the Arizona Department of Economic Security (DES) is the State agency responsible for providing official population and employment estimates and projections for Arizona and its counties. DES has annual employment statistics until year 2005. However, DES does not forecast special populations (group quarters, transient and seasonal) and housing units for these populations.

In order to extend its five-year forecasts through year 2040, MAGTPO estimated region control totals based on the resident population projections provided by DES. County projection of total employment for the region was calculated as a percentage of DES population estimates. After total employment forecasts for the county were generated, employment figures were then broken into eight main DES sectors used in DES reports (Manufacturing; Mining; Construction; Transportation, Communications and Public Utilities (TCPU); Trade; Finance, Insurance and Real Estate (FIRE); Services; and, Government) and then converted into five MAG sectors (Retail, Office, Industrial, Government and Others). Conversion into MAG sectors was necessary since these sectors were required in the DRAM/EMPAL programs. A separate fixed estimate for military employment was added to the government employment total.

MAG also estimated a separate special population figure, which included the group quarters population (military, jails, colleges, nursing homes, others), transient population (nonresident population staying in the region for at most two weeks) and seasonal population (nonresident population staying in the region for more than two weeks). The total county population was the sum of the resident population and the special population.

MAG also projected the total number of households or units (jail cells, boarding/dormitory rooms, hotel/motel rooms, mobile homes, recreational vehicles, other seasonal homes). The total number was then divided by five to arrive at the household quintile estimate.

These estimates were then used as inputs to the DRAM/EMPAL land use forecasting model fitting programs.

4.3 Review of MAG Procedures

To facilitate discussion of procedures implemented by MAGTPO, a series of Excel spreadsheets is shown in Figure 4-1. MAGTPO derived its employment total as a percentage of DES projection of the Maricopa County resident population (see shaded cells). Assuming that the DES figures are valid, validity of employment estimates depends on the validity of the projected wage and salary equivalence ratios (W&S ratio). These ratios were based on US national projections by the US Bureau of Economic Analysis. The projected ratios suggested that employment' share would steadily peak at year 2005 and then slowly decline thereafter.

To validate MAGTPO's wage and equivalence ratio, MAGTPO's 1995 and 2000 W&S ratios were compared with DES figures. For year 1995, DES Maricopa County employment forecast yielded a ratio of 0.444 - insignificantly lower than MAGTPO ratio of 0.45. For year 2000, DES figures produced a ratio of 0.455 which is less than MAGTPO's projection of 0.47. Note that MAGTPO's W&S ratio assumptions consistently generated higher employment totals than estimated by DES for years 1995 and 2000.

Figure 4-1

MAG Derivation of Control Totals from DES Population Projections

	1990	1995	2000	2005	2010	2015	2020	2025	2030	2035	2040
DES Population	2,130,400	2,415,411	2,716,449	3,032,994	3,364,885	3,718,770	4,092,176	4,476,898	4,867,999	5,266,536	5,677,127
W&S Equiv Ratio	0.45	0.45	0.47	0.48	0.47	0.45	0.44	0.43	0.42	0.41	0.41
Total Employment	964,200	1,098,390	1,279,936	1,455,078	1,569,560	1,678,849	1,786,041	1,920,378	2,051,631	2,180,097	2,307,484
Employment by DES Sector											
Manufacturing	139,600	132,440	146,823	157,292	162,096	164,964	167,884	173,963	180,606	188,012	196,312
Mining	800	802	803	758	671	570	474	395	330	283	252
Construction	56,300	64,244	75,014	84,865	92,353	98,954	105,426	113,489	121,352	129,030	136,624
TCPU	58,200	61,309	69,669	76,657	81,194	84,860	88,480	93,588	98,745	104,006	109,450
Trade	242,800	272,131	315,982	355,479	384,788	410,317	435,372	467,135	498,275	528,889	559,390
F.I.R.E.	75,800	82,589	93,167	101,718	106,899	110,896	114,860	120,818	126,928	133,278	139,966
Services	269,500	325,831	393,656	460,220	516,140	567,747	618,165	676,847	732,873	786,026	836,952
Government	121,200	159,045	184,818	208,082	225,407	240,526	255,361	274,119	292,495	310,542	328,503
Military (*)	9,000	9,000	9,000	9,000	9,000	9,000	9,000	9,000	9,000	9,000	9,000
Trended Out Employment Sector Shares (See ESCTREND.XLS)											
Manufactoring	0.1448	0.1206	0.1147	0.1088	0.1033	0.0983	0.0940	0.0906	0.0880	0.0862	0.0851
Mining	0.0008	0.0007	0.0006	0.0005	0.0004	0.0003	0.0003	0.0002	0.0002	0.0001	0.0001
Construction	0.0584	0.0585	0.0586	0.0587	0.0588	0.0589	0.0590	0.0591	0.0591	0.0592	0.0592
TCPU	0.0604	0.0558	0.0544	0.0530	0.0517	0.0505	0.0495	0.0487	0.0481	0.0477	0.0474
Trade	0.2518	0.2478	0.2469	0.2460	0.2452	0.2444	0.2438	0.2433	0.2429	0.2426	0.2424
FIRE	0.0786	0.0752	0.0728	0.0704	0.0681	0.0661	0.0643	0.0629	0.0619	0.0611	0.0607
Services	0.2795	0.2966	0.3076	0.3185	0.3288	0.3382	0.3461	0.3525	0.3572	0.3605	0.3627
Government	0.1257	0.1448	0.1444	0.1440	0.1436	0.1433	0.1430	0.1427	0.1426	0.1424	0.1424

DES/MAG Employment Sector Conversion

	Retail	Office	Industrial	Gov't	Other
Manufacturing	0.0000	0.0946	0.8932	0.0122	0.0000
Mining	0.0000	0.1071	0.5963	0.0000	0.2966
Construction	0.0000	0.0466	0.2676	0.0000	0.6859
TCPU	0.0875	0.1264	0.2275	0.0790	0.4796
Trade	0.4973	0.1069	0.1530	0.0741	0.1688
F.I.R.E.	0.1446	0.7885	0.0000	0.0000	0.0669
Services	0.1081	0.1565	0.1225	0.2770	0.3359
Government	0.0000	0.0000	0.0000	1.0000	0.0000

Employment by MAG Sector

	1990	1995	2000	2005	2010	2015	2020	2025	2030	2035	2040
Retail	165,921	187,849	219,248	247,932	269,697	288,869	307,668	331,115	353,991	376,339	398,454
Office	151,192	168,584	195,150	218,863	236,203	251,194	265,916	284,783	303,340	321,659	339,990
Industrial	223,631	231,458	264,108	291,855	310,464	325,795	340,911	361,662	382,462	403,496	425,067
Government	229,132	284,910	333,553	378,860	414,261	445,889	476,866	514,708	551,397	586,940	621,797
Other	203,350	234,617	276,907	316,599	347,962	376,130	403,706	437,134	469,466	500,688	531,200
Total	973,226	1,107,418	1,288,966	1,454,109	1,578,587	1,687,877	1,795,067	1,929,402	2,060,656	2,189,122	2,316,508
Special Populations											
ResHHPop	2,097,090	2,377,749	2,673,726	2,985,100	3,311,448	3,659,241	4,026,477	4,405,002	4,789,998	5,182,620	5,587,545
ResGQPop	33,310	37,662	42,723	47,894	53,437	59,529	65,699	71,896	78,001	83,916	89,582
Total ResPop	2,130,400	2,415,411	2,716,449	3,032,994	3,364,885	3,718,770	4,092,176	4,476,898	4,867,999	5,266,536	5,677,127
TranPop	36,718	41,790	48,071	53,633	59,803	66,419	72,799				
SeasPop	46,577	52,907	60,622	67,617	75,175	83,308	91,094				
TotalPop	2,213,695	2,510,108	2,825,112	3,154,244	3,499,863	3,868,497	4,256,069				

Households

	1990	1995	2000	2005	2010	2015	2020
AvgHHOcc	2.59	2.59	2.59	2.58	2.58	2.58	2.57
ResHHOcc	810,219	918,050	1,032,327	1,157,016	1,283,507	1,418,310	1,566,723
GQHHOcc	6,810	7,759	8,814	9,876	10,999	12,232	13,469
TranHH	29,322	33,333	38,300	42,756	47,678	52,966	58,050
SeasHH	31,046	35,274	40,411	45,076	50,129	55,540	60,742
TotalHH	877,397	994,416	1,119,852	1,254,724	1,392,313	1,539,048	1,698,984
ResHHQuintiles		183,610	206,465	231,403	256,701	283,662	313,345

In order to check MAGTPO's employment projection for year 2005, a comparison was made between the difference in DES projections for years 2000 and 2002, and the difference between DES year 2002 estimate and MAGTPO year 2005 projection. If these two differences were comparable, then the year 2005 employment estimate by MAGTPO could be considered attainable. The two-year and three-year differences are given in Table 4-1. DES estimated employment for year 2002 to reach only 1.31 million whereas MAG estimated employment levels at 1.45 million by 2005. Comparison of the last two columns of Table 4-1 revealed that the growths projected by MAGTPO were too optimistic: the three-year increase (10.6% overall) was about twice as much as the two-year DES increase (5.6% overall) in all sectors except mining.

Table 4-1

Comparison of DES Year 2002 and MAG Year 2005 Employment Estimates by Sector

Sector	Employment Estimates*			Difference from 2000 to 2002	Difference from 2002 to 2005
	DES Year 2000	DES Year 2002	MAG Year 2005		
Manufacturing	141,904	147,727	157,292	5,823	9,565
Mining	776	774	758	(2)	(16)
Construction	72,501	75,962	84,865	3,461	8,903
TCPU	67,335	70,330	76,657	2,995	6,327
Trade	305,395	322,097	355,479	16,702	33,382
FIRE	90,045	93,904	101,718	3,859	7,814
Services	380,466	407,326	460,220	26,860	52,894
Government	178,625	188,438	208,082	9,813	19,644
Total	1,237,050	1,306,557	1,445,071	69,507	138,514

*Source: Arizona Department of Economic Security, Research Administration, February 1993. Maricopa Association of Governments Transportation and Planning Office.

The MAGTPO's employment projections for years 1995 to 2005 were found to be higher than the estimates made by DES. Although the projected ratios were consistent with the changing age distribution tending towards a more senior adult population, AMPG recommends that the wage and salary equivalence ratio projections (see Figure 4-1) be made consistent with the DES projections for years 2000 and 2002. This has the effect of lowering the W&S ratio in the years 2000 and 2005. Table 4-2 summarizes AMPG's proposed adjustments.

Table 4-2

AMPG Modified Wage and Salary Equivalence Ratio

1995	2000	2005	2010	2015	2020	2025	2030	2035	2040
0.45	0.46	0.465	0.47	0.46	0.45	0.44	0.43	0.42	0.41

Breakdown into DES Sectors

The second major task involved the breakdown of the employment totals into DES industry sectors. MAGTPO used the 1995 and 2000 DES projections as the basis for projecting the employment percentage share by eight sectors up to year 2040. It was assumed that any increase or decrease in the employment share from 1995 to 2000 for each sector would also apply to succeeding five-year periods. Although this assumption seemed reasonable, the formula for deriving the trended out employment sector shares used was inadequate, particularly for the Mining sector where employment was estimated to be only 252 by the year 2040. MAGTPO calculated sector employment shares for a particular five-year period by expressing it as the preceding five-year employment share plus the difference in employment share for the preceding two year-year periods weighted by a linearly declining constant. The weights used by MAGTPO resulted in a decrease in the mining employment share. To minimize this drop in employment, AMPG recommends that a smoother exponential function be used as weights which would provide more stable sector share estimates every five years. The revised formula for generating employment market share is:

$$\%\ share_i = \%\ share_{i-1} + (\%\ share_{i-1} - \%\ share_{i-2})\ x\ \exp(-0.1i),$$

where $i = 1,\ldots,8$ indexes the 5-year periods starting on year 2005. Applying this formula, the estimated employment share by sector is given in Table 4-3.

Table 4-3
AMPG Revised Employment Sector Shares

Sector	2005	2010	2015	2020	2025	2030	2035	2040
Manufacturing	0.1094	0.1051	0.1018	0.0997	0.0984	0.0977	0.0973	0.0971
Mining	0.0005	0.0005	0.0004	0.0004	0.0003	0.0003	0.0003	0.0003
Construction	0.0587	0.0588	0.0589	0.0589	0.0589	0.0590	0.0590	0.0590
TCPU	0.0532	0.0522	0.0514	0.0509	0.0506	0.0504	0.0503	0.0503
Trade	0.2461	0.2454	0.2449	0.2446	0.2444	0.2443	0.2443	0.2442
FIRE	0.0706	0.0688	0.0675	0.0666	0.0661	0.0658	0.0657	0.0656
Services	0.3174	0.3255	0.3315	0.3355	0.3380	0.3393	0.3400	0.3403
Government	0.1440	0.1437	0.1435	0.1434	0.1433	0.1432	0.1432	0.1432

Applying both the modified W&S ratio and employment sector shares, the revised employment totals are given in Table 4-4. The employment share by the mining sector was stabilized around 700 throughout the projection period. Employment share by the Services sector for year 2005 would increase by 40,367 from DES year 2002 projection. The Trade sector share would expand by 24,956 more workers from DES year 2002 projection. These sector growth projections are still realistic considering the yearly projections of DES.

Table 4-4
AMPG Revised Employment by DES Sector

Sector	2000	2005	2010	2015	2020	2025	2030	2035	2040
Manufacturing	143,340	154,298	166,152	174,213	183,565	193,782	204,418	215,222	226,104
Mining	784	754	725	688	672	673	689	714	745
Construction	73,235	82,809	92,998	100,703	108,487	116,101	123,404	130,419	137,247
TCPU	68,016	75,000	82,479	87,913	93,699	99,621	105,507	111,304	117,037
Trade	308,485	347,053	388,138	419,005	450,458	481,469	511,406	540,289	568,488
FIRE	90,956	99,595	108,868	115,503	122,712	130,209	137,751	145,240	152,683
Services	384,316	447,693	514,811	567,096	617,867	665,732	710,236	751,981	792,001
Government	180,432	203,135	227,315	245,499	264,005	282,230	299,808	316,756	333,295
Total	1,249,567	1,425,507	1,581,496	1,710,634	1,841,479	1,969,835	2,093,240	2,211,945	2,327,622

Employment by Military

Military employment was fixed at 9,000 workers throughout the projection period. This estimate was consistent with the 1990 Census figure of 8,734. No expansion or down-sizing was expected to the existing Luke Air Force Base in Maricopa County. However, Williams Air Force Base was slated for demobilization before 1995. Thus, it is possible that the 1990 Census figure includes employment at Williams which no longer applies. As such, the constant figure assumption should be revisited.

Conversion into MAG Sectors

The next step in the MAGTPO procedure was the conversion of employment shares by DES sector into MAG employment sector shares. MAGTPO used a conversion matrix based on ESG 1990 adjusted land use matrix (see Table V-4 of the ESG Final Report). MAGTPO did further adjustment by separating the Education component from the Services category and reclassifying it under the Government sector.

The MAGTPO conversion matrix is found in Figure 4-1. For example, the Manufacturing sector share was divided into three MAG sectors (Office, Industrial and Government) with allocation proportion defined in the first row of the matrix (9%, 89% and 1%). Estimated employment totals by MAG sector are shown on the second page of Figure 4-1.

The conversion procedure implemented by MAGTPO was methodologically sound. However, employment county totals may not necessarily match the initial county estimates because of round-off errors. AMPG suggests that a further adjustment be made to ensure that employment county totals remained fixed in all tables. AMPG further recommends that the conversion factors be updated as soon as updated statistics are obtained from DES.

Estimation of Special Population

MAGTPO used special population estimates other than those provided by ESG. For comparison, Table 4-5 shows MAG and ESG estimates for group quarters, transient and seasonal population. MAG's special population estimates were consistent with the 1985 population estimates provided by DES. Estimates of ESG group quarters population estimates were lower than MAG estimates except in the last two years: 2035 and 2040. MAG transient population estimates were between ESG average and peak estimates until year 2000; therafter, MAG estimates were still higher than ESG peak estimates. There was a wide discrepancy in the seasonal population estimates. This was probably due to differences in the definition of seasonal population. ESG included seasonal population living in mobile homes, recreation vehicles and other housing. Unfortunately, there was no additional documentation available on MAG's projections.

Table 4-5
MAG and ESG Projections for Special Population

	Group Quarters[1]		Transient[2]			Seasonal[3]	
Year	MAG	ESG	MAG	ESG (average)	ESG (peak)	MAG	ESG
1985	28,371[4]	no estimate	34,448[4]	no estimate	no estimate	45,045[4]	164,965
1990	33,310	33,310	36,718	29,941	40,792	46,577	164,965
1995	37,662	35,823	41,790	32,432	44,185	52,907	178,570
2000	42,723	39,640	48,041	35,336	48,142	60,622	197,570
2005	47,894	43,918	53,633	37,810	51,513	67,617	216,439
2010	53,437	48,965	59,803	39,713	54,104	75,175	239,202
2015	59,529	54,851	66,419	41,227	56,167	83,308	268,569
2020	65,699	61,049	72,799	42,765	58,263	91,094	299,927
2025	71,896	68,808	no estimate	44,417	60,513	no estimate	329,950
2030	78,001	77,530	no estimate	46,044	62,730	no estimate	363,251
2035	83,916	85,203	no estimate	47,607	64,860	no estimate	394,608
2040	89,582	91,419	no estimate	49,080	66,866	no estimate	426,022

[1] Source: Maricopa Association of Governments Transportation and Planning Office. Table IV-1, Economic Strategies Group Final Report.
[2] Source: Maricopa Association of Governments Transportation and Planning Office. Table IV-4, Economic Strategies Group Final Report.
[3] Source: Maricopa Association of Governments Transportation and Planning Office. Table IV-10, Economic Strategies Group Final Report.
[4] Source: 1985 Special Census for Maricopa County.

Since the source of the discrepancy, particularly on the seasonal population, could not be identified , AMPG recommends that a sensitivity be applied to assess the effects of these two estimates on the DRAM/EMPAL models.

Estimation of Number of Households

The projected numbers of resident households were obtained by dividing resident population estimates by average household occupancy values. The assumed average household occupancy seemed to be reasonable. Census estimated around 808,000 occupied housing units in 1990 in Maricopa County; the MAGTPO estimate was 810,000 households.

It also appeared that MAGTPO derived the number of housing units for special population by dividing the special population estimate by a fixed constant. MAGTPO used factors of 4.85, 1.254 and 1.5 for the group quarters, transient and seasonal populations, respectively. In the ESG Final Report, the average occupancy rate for the transient population ranged from 1.0 to 2.0 depending on the group category (leisure, group, business and others). For the years 1990 and 2040, the weighted average occupancy rates were 1.52 and 1.41, respectively. For the seasonal population, ESG used a factor of 2.0 per housing unit. MAGTPO's factors were consistently below ESG's numbers. If applied to a population estimate, MAGTPO's procedure would result in more housing units than would have been derived if ESG's occupancy rate factors were used.

Table 4-6 provides a comparison of MAG and ESG estimates of housing units for special population. MAG's 1990 estimates were consistent with the 1985 housing figures provided by DES. There were huge differences in the estimates, particularly for seasonal housing units. One possible source of the discrepancies was the use of different occupancy rates. Another source could be due to differences in the definition of occupied housing units. The needs of the DRAM/EMPAL programs should be further investigated to determine which estimates should be used. As with the seasonal population estimates, AMPG recommends that a sensitivity analysis be applied to assess the effects of these two estimates on the DRAM/EMPAL models.

Table 4-6

MAG and ESG Projections for Special Population Occupied Housing Units

	Group Quarters[1]		Transient[2]			Seasonal[3]	
Year	MAG	ESG	MAG	ESG (average)	ESG (peak)	MAG	ESG
1985	5,674[4]	no estimate	27,558[4]	no estimate	no estimate	33,367[4]	no estimate
1990	6,810	no estimate	29,322	21,485	29,271	31,046	82,482
1995	7,759	no estimate	33,333	23,485	31,996	35,274	89,285
2000	8,814	no estimate	38,300	25,974	35,387	40,411	98,632
2005	9,876	no estimate	42,756	28,099	38,283	45,076	108,220
2010	10,999	no estimate	47,678	29,707	40,472	50,129	119,601
2015	12,232	no estimate	52,966	31,014	42,253	55,540	134,284
2020	13,469	no estimate	58,050	32,346	44,069	60,742	149,964
2025	no estimate	no estimate	no estimate	33,889	46,170	no estimate	164,975
2030	no estimate	no estimate	no estimate	35,412	48,245	no estimate	181,626
2035	no estimate	no estimate	no estimate	36,877	50,241	no estimate	197,304
2040	no estimate	no estimate	no estimate	38,259	52,124	no estimate	213,011

[1] Source: Maricopa Association of Governments Transportation and Planning Office. No estimates were provided by ESG.
[2] Source: Maricopa Association of Governments Transportation and Planning Office. Table IV-4, Economic Strategies Group Final Report.
[3] Source: Maricopa Association of Governments Transportation and Planning Office. ESG figures were derived by dividing the seasonal population estimate from Table 4-5 by two.
[4] Source: 1985 Special Census for Maricopa County.

4.4 Review of ESG Estimation of Special Population

The Economic Strategies Group (ESG) performed five-year county-level projections of special population (group quarters, transient and seasonal population) from 1995 through 2040. AMPG's review of ESG's forecasts was based on the estimation procedures documented in the ESG Final Report.

Estimation of Group Quarters Population

Group quarters population consisted of individuals who were residing in military bases, jails and prisons, colleges and universities, nursing homes and other group quarters (including mental and juvenile institutions, emergency shelters, etc.). ESG estimated the population in each category and then summed all of the estimates to obtain the group quarters population estimate.

The military group quarters population was based on the 1990 Census and was fixed throughout the projection period. ESG removed 314 individuals to reflect the closing of the Williams Air Force Base by 1995.

The jail and prison population was expressed as a fixed percentage (0.399%) of the county resident population throughout the projection period. This fixed share assumption, however, did not take into account the changing age distribution patterns over the 45-year period. As was done for student and nursing home populations, a more reliable estimate would have been obtained by using a fixed percentage of the adult population.

The student population housed in group quarters was derived as a fixed percentage (3.0785%) of the college-aged county residents (18 to 22 years). However, it seemed unlikely that estimates for the years 1993 and 1995 would be lower than the 1990 figure; ESG estimated student population at 5,256, 4,242 and 4,868 for the years 1990, 1993 and 1995, respectively.

Similarly, the nursing home population was based on a fixed percentage (3.26%) of the senior citizen population (65 years or older). The other quarters population was calculated using a fixed percentage of 0.45% of the resident population.

The estimation approach for the group quarters components was based on a fixed share of the total population or a specific sub-group. Validity of the ESG projections rests on these constant shares holding throughout the 45-year span. AMPG did not find any major flaw in ESG's assumptions. Since there were slight discrepancies between total population estimates used by ESG and MAG, AMPG recommends that the group quarters population be updated using DES population projections currently being used by MAG.

ESG did not provide housing unit estimates associated with the group quarters population. AMPG further suggests that occupancy rates be estimated for all group quarters categories and applied to the population projections to generate housing unit forecasts. The baseline occupancy rates may be obtained from the key persons contacted by ESG for its baseline data.

Transient Population

Transient population was comprised of hotel and motel guests divided into four categories depending on the travel purpose (leisure, business, group related travel and others).

The leisure population was projected using a fixed percentage of total population from each state or country. As expected, visitors from California were the largest group. The second most largest group were from other southern states. It was surprising, however, that ESG did not provide separate estimates for these southern states.

Although the ESG estimation procedure may apply to domestic travellers, different economic mechanisms may be responsible for the influx of tourists from other countries. In particular, money exchange rates are more likely to impact foreign travel. In the ESG report, over 70 percent of foreign leisure visitors were from Mexico. Since the population of Mexico was projected to increase at a much faster rate than the United States population, foreign visitors accounted for over 15 percent of total visitors by the year 2040 (see Table 4-7). This unusually high foreign share of visitors would be plausible if the economic situation in Mexico would improve in conjunction with Mexico's population growth. On the other hand, visitors from Japan and Germany were projected to decline because of the downward trend in the population of the two countries. Thus, ESG's procedure fails to account for other important economic factors influencing foreign travel.

Table 4-7
ESG Projected Domestic and Foreign Leisure Visitors

Source	1990	1995	2000	2005	2010	2015	2020	2025	2030	2035	2040
Domestic	88.0%	87.6%	87.2%	86.8%	86.4%	86.1%	85.7%	85.3%	84.9%	84.6%	84.3%
Foreign	12.0%	12.4%	12.8%	13.2%	13.6%	13.9%	14.3%	14.7%	15.1%	15.4%	15.7%
Total	6,373	6,796	7,179	7,428	7,566	7,512	7,458	7,404	7,351	7,297	7,243

If foreign visitors are not expected to increase their share of the leisure population, an alternative approach for estimating the number foreign leisure visitors is to express the number as a fixed percentage of the domestic visitors (e.g., 12 percent). Although the recommended approach still does not take into account influential economic factors, this procedure will lead to lower estimates of leisure visitors.

The group visitors population was derived using a fixed percentage of employment projections from other states. Group visitors from foreign countries were not included. Similarly, the business visitors population was expressed as a fixed percentage of employment in the Maricopa County. Foreign group and business visitors may be accounted for by taking a percentage of the domestic group and business visitors. After the first three visitor populations were projected, the other visitors population was projected using a fixed percentage (9.77%) of the sum of the three visitor population groups.

Visitors population at peak season was estimated by inflating the average projections by 1.3624 (derived by ESG from the Phoenix & Valley of the Sun Convention & Visitors Bureau).

ESG also projected available and occupied rooms (housing units) throughout the projection period. The number of occupied rooms can be used for MAGTPO's projection.

Seasonal Population

ESG estimated the seasonal population residing in mobile homes, recreational vehicles and other housing. Seasonal population in mobile homes was estimated using a fixed percentage of the US (by state) and southwestern Canadian residents aged 65 and up. Since the Canadian population was projected to grow at a much faster rate than the US population, Canadian's share increase from 17 percent in 1990 to 25 percent in year 2040 (see Table 4-8). As was suggested for the leisure visitor group, a more conservative estimate of the foreign seasonal dwellers could be derived by taking a fixed percentage (17%) of the US seasonal dwellers.

Table 4-8
ESG Projected Domestic and Foreign Seasonal Visitors in Mobile Homes

Source	1990	1995	2000	2005	2010	2015	2020	2025	2030	2035	2040
Domestic	83.0%	82.7%	82.0%	81.4%	80.8%	80.6%	80.2%	79.3%	76.4%	75.7%	75.1%
Foreign	17.0%	17.3%	18.0%	18.6%	19.2%	19.4%	19.8%	20.7%	23.6%	24.3%	24.9%
Total	44,185	45,364	46,714	48,460	52,610	61,132	69,811	76,022	84,453	90,780	97,207

The seasonal population residing in recreational vehicles was projected with the same methodology used for mobile home seasonal dwellers. Table 4-9 presents the projected seasonal population in recreational vehicles from the US and Canada. Comparison of Table 4-9 with Table 4-8 suggested that the distribution of seasonal dwellers from the different states and Canada was the same for both mobile home and seasonal dwellers. Further examination of Tables IV-12 and Tables IV-14 of the ESG Final Report confirmed that only one distribution was used for both mobile homes and recreational vehicles. Unfortunately, AMPG did not have access to the 1992 Winter Resident Study conducted by Arizona State University and thus, AMPG could not verify this assumption. It seemed unlikely, however, that the distribution of place of residence of the seasonal travellers would exactly be identical for both mobile homes and recreational vehicles. AMPG recommends that this assumption be further investigated.

Table 4-9
ESG Projected Domestic and Foreign Seasonal Visitors in Recreational Vehicles

Source	1990	1995	2000	2005	2010	2015	2020	2025	2030	2035	2040
Domestic	83.0%	82.7%	82.0%	81.4%	80.8%	80.6%	80.2%	79.3%	76.4%	75.7%	75.1%
Foreign	17.0%	17.3%	18.0%	18.6%	19.2%	19.4%	19.8%	20.7%	23.6%	24.3%	24.9%
Total	50,072	51,407	52,937	54,914	59,617	69,275	9,110	86,148	95,701	102,871	110,154

The number of mobile homes and recreational vehicles allocated for seasonal population was derived by dividing the projected population by two (i.e., two persons per unit).

Starting from total housing units projected by MAG, ESG projected the seasonal population quartered in other housing by multiplying available units by two. Available units were derived by taking 5.905% of the total housing units and subtracting mobile home units allocated for seasonal population. The factor of 5.905% was the 1990 percent share of vacant units held for seasonal, migrant or other uses. It seemed reasonable to decrease this value to isolate the share of vacant units primarily used during peak seasons. Since the source of MAG's projected total housing units could not be ascertained, AMPG recommends that the MAG projected total housing units and the fixed percentage of 5.905% be reviewed to ensure consistency with DES projections.

Conclusions

After examination of procedures implemented by ESG to estimate county-level special population, AMPG's conclusions and recommendations were as follows:

- ESG estimation of the group quarters population was acceptable. However, AMPG still recommends the use of DES projected population estimates to update the group quarters population.

- AMPG recommends a further study on the occupancy rates for the group quarters population to estimate housing units occupied by the group quarters population.

- ESG estimation of the domestic transient population was sufficient. Projections on foreign visitors relied heavily on population estimates of selected countries, particularly Mexico, and did not consider other important economic factors. Alternative methods should be devised. If foreign visitors were not expected to account for more share of the total transient population, then an alternative method would be to express foreign visitors as a function of domestic visitors. This alternative approach would tend to lower estimates for the leisure visitors group.

- Group and business visitors from other countries were not accounted for in the ESG estimation. Thus, possible overestimation in the leisure visitors group might have been offset by underestimation in the group and business visitors group.

- Estimation of the seasonal population requires further investigation. ESG used an identical percent share of visitors from different states and Canada for estimating mobile home and recreational vehicle dwellers. Seasonal population in other housing was dependent on MAG's projection of total housing units, assumption of a fixed percentage of total housing to arrive at available housing and assumption of 2.0 individuals per unit. The last assumption might have created the unusually large seasonal population in other housing.

Overall, AMPG recommends that the base population used in most projections be reviewed and updated using DES projections.

5
Develop Floor Area Ratio and Square Feet per Employee Data

MAG relies heavily on DRAM/EMPAL for generating socioeconomic forecasts throughout the region for regional analysis zones (RAZs), based on forecast control totals applicable to the entire region. EMPAL is the component of the model responsible for generating employment forecasts for five categories of workers: (1) office workers, (2) industrial workers, (3) retail workers, (4) public (government) workers, and (5) other employment. Through the concept of constraints, the model takes into consideration the availability of land for the various types of land use development associated with growth in these employment sectors. Consequently, an important element of the forecasting process deals with tracking the absorption of land in RAZs associated with growth.

Two mechanisms are involved:

- Square footage per employee ratios to associate building space with employment;
- Floor area ratios to associate land absorption with building sizes

Data pertinent to these two ratios is the subject of this section.

5.1 Task Objectives

Initially, this task called for a review of current MAG procedures establishing square foot per employee ratios and floor area ratios. Clearly these ratios are key to successful forecasting: they provide the basis for translating socioeconomic projections of housing and employment from DRAM/EMPAL into growth in land use and development. DRAM/EMPAL needs this information so RAZ *constraints*[1] can be properly set.

Through the course of the project it *also* became evident that the development of the Subarea Allocation Model (SAM) was one of the project's *most important objectives*. The thrust of the SAM is to allocate DRAM/EMPAL forecasts for RAZs to a finely grained microscale geography called the Minimal Analysis Zone (MAZ) system. Once allocated, MAG would then have the ability to aggregate socioeconomic forecast information to whatever level of geography is desired, whether it be TAZs, municipal boundaries, school districts, water districts, or whatever.

The SAM too will be driven by land absorption rates, but will require other information as well. Consequently, this chapter will address the data required to make the SAM operational.

5.1.1 The Subarea Allocation Model

A full discussion of the approach to the SAM is given in Chapter 8 and readers might benefit from a review of the material there. Here is a brief overview of the basic approach because an understanding of the SAM helps us identify now what kinds of land absorption data is needed.

The fundamental concept behind the SAM is to allocate housing and employment *growth* projections generated by DRAM/EMPAL on an RAZ basis to a rasterized geography of grid cells (the MAZ) covering Maricopa County.

An *existing land use* cover which specifies the location of developed land, its use and density, reports for the base year the distribution of households and employment (and also identifies lands which are available for development, and therefore are candidates to absorb growth). For the *base year*, MAG's existing land use cover (for 1994) will serve this purpose. This cover is shown in Figure 5-1.

The *general plan cover* identifies how lands available to absorb growth can be used. The SAM will select among candidate lands that are appropriately classified which ones will be developed to accommodate growth implied in the DRAM/EMPAL projections. The combination of existing land use in the base year, and those allocated development for the future year, will yield a *new* land use cover which provides a basis for subsequent future years.

The SAM then will be generating a series of land use covers meant to describe then-existing land use for the forecast year in question. The idea will be to develop these databases in sequence, so the existing land use cover for the year 2000, say, becomes the basis for generating the existing land use cover for the year 2005. These forecast land use covers,

---

[1] The availability of land to accommodate growth in RAZs is a clear constraint that DRAM/EMPAL will obey. The program, however, must be told how much land is available in RAZs to accommodate growth.

Figure 5-1a
Existing Land Use Cover for Maricopa County

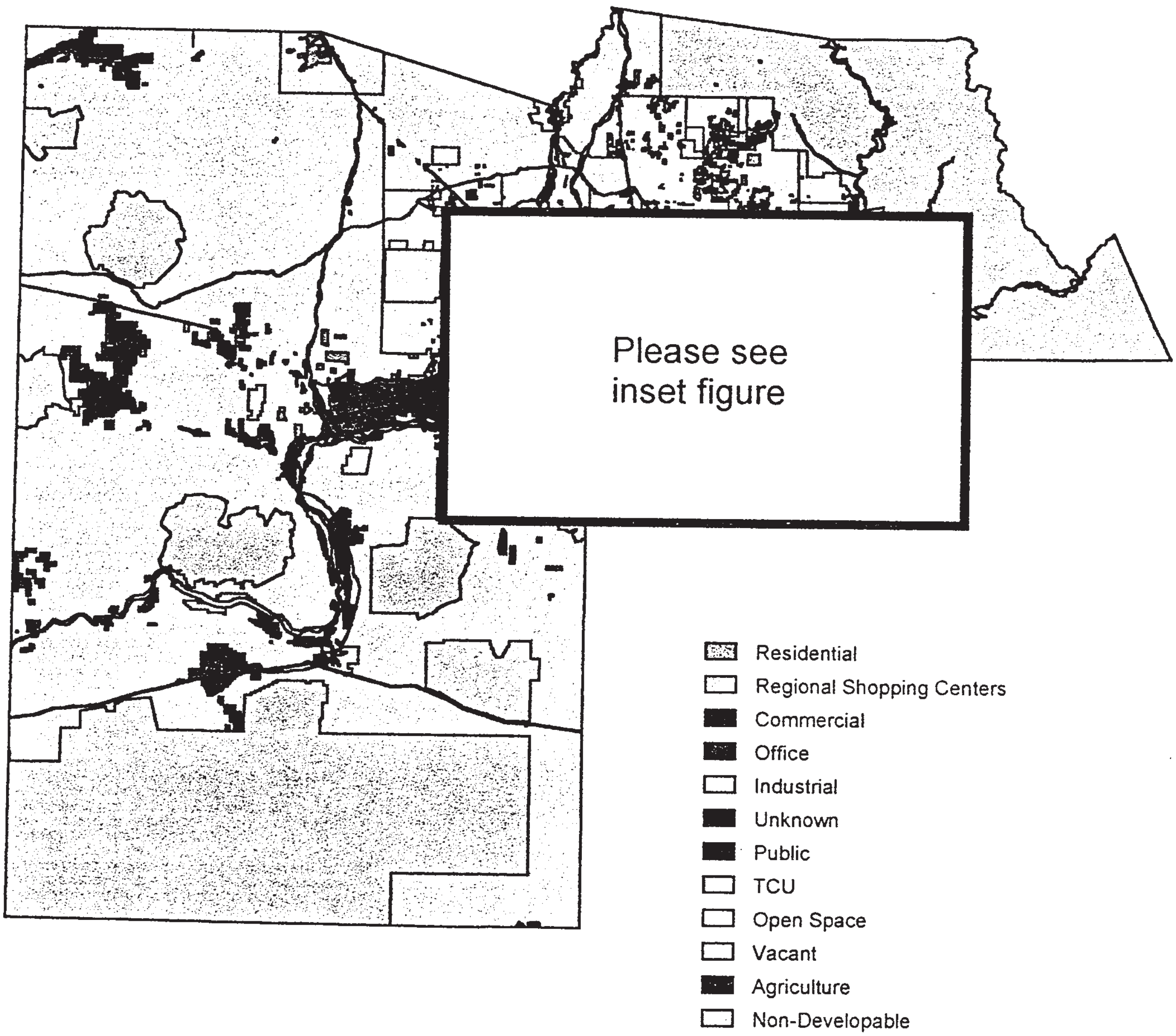
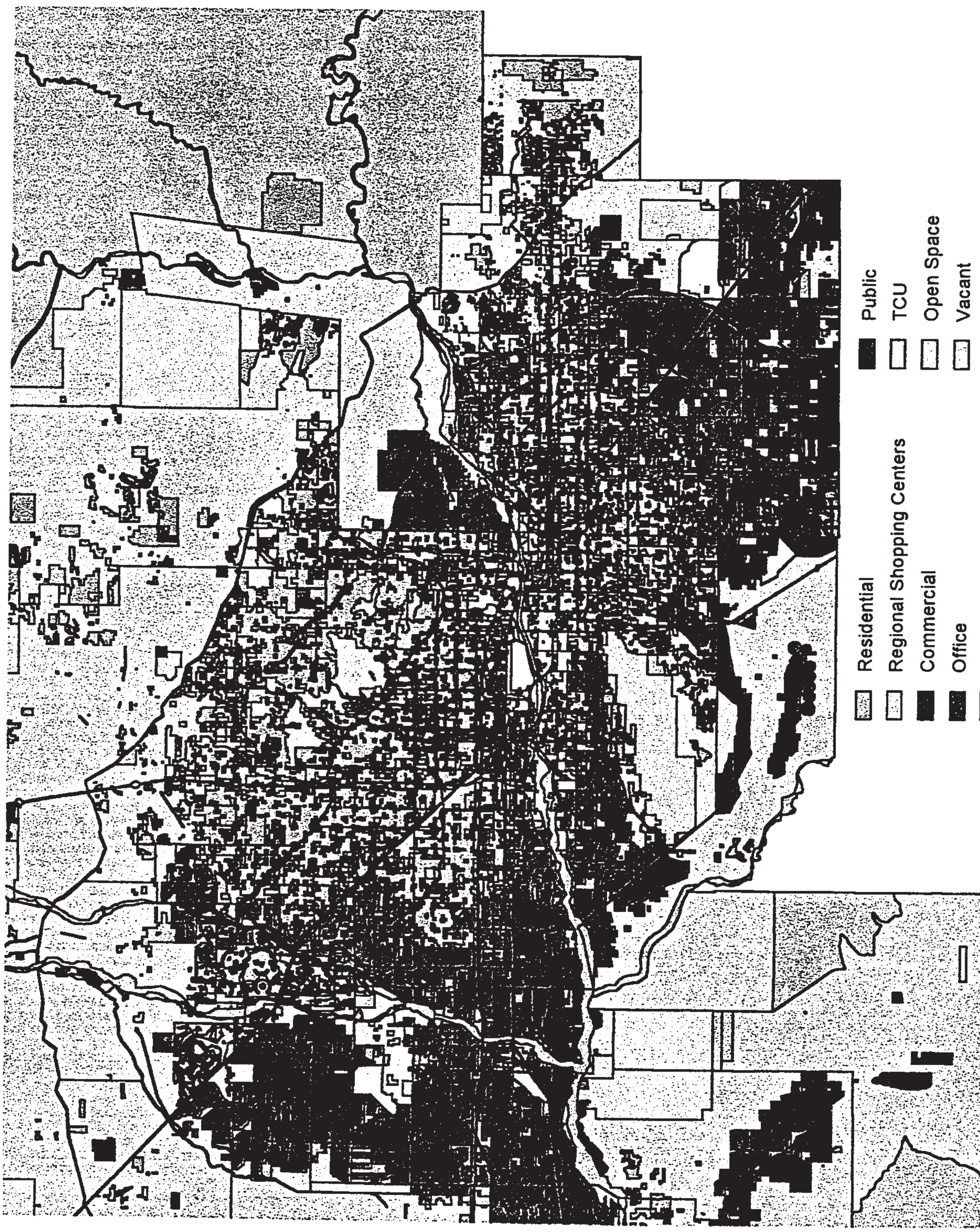


Figure 5-1b
Existing Land Use Cover for Maricopa County (detail)



then, represent a *simulation* of growth and development in the region throughout the forecast period. The general plan coverage, which plays a key role in identifying the future use of lands, is shown in Figure 5-2.

The important concept behind the SAM is that it allocates growth, not overall household and employment totals for the forecast year. If it can establish an existing land use cover for 1994 which accurately portrays the distribution of housing and employment (among the 13,000 polygons covering Maricopa County), this will minimize error in future projections by concentrating only on changes in RAZs.

5.1.2 The Simulation Mechanism

The simulation mechanism to be implemented in the SAM requires these kinds of issues to be addressed:

- Growth in Households

  - Which general plan land use categories are eligible to absorb housing growth needs to be defined. Of course, these will be the residential categories, however agricultural lands and otherwise undeclared lands should also be considered to be eligible.

  - For each general plan category, the density levels at which housing will be built needs to be defined (that is, the land absorption rate).

  There will be other issues that will have to be addressed, for example (1) the relative attraction of different residential types to different income levels associated with household growth projected by DRAM/EMPAL, (2) aging of the housing stock,[2] and (3) locational preferences for housing development.

- Growth in Employment

  - For each of the five employment classes, general plan land use designations that are candidates needs to be identified for their development.

  - Also, overall gross land absorption rates associated with commercial developments of different densities needs to be defined.

5.2 Existing Information on Density Levels

Prior research at MAG has generated information on prevailing densities of development.

---

[2] MAG wants to add the age of housing stock to its basic variable set.

Table 5-1 reports prevailing land use densities for general categories of residential development. As will be noted shortly, however, the general plan cover itself defines residential lands by density level, and therefore those shown in Table 5-1 will not be needed except in the absence of general plan information.

Table 5-1
General Residential Densities

Residential Classification	Low Bound	High Bound
Very Low	0.0	0.2
Low Rural	0.2	0.5
Rural	0.5	1.0
Low Density	1.0	2.5
Medium Low Density	2.5	4.0
Low Medium Density	4.0	6.0
High Medium Density	6.0	8.0
Medium High Density	8.0	10.0
High Density	10.0	15.0
Very High Density	15.0	25.0
Highest Density	25.0	—

Table 5-2 reports prevailing square-foot-per-employee ratios (of building space) associated with various types of employment. However, without similar floor-area-ratio information for each land use to tie to these ratios, the land absorption rates cannot be developed for the SAM. Also, the general land use categories named in Table 5-2 to the specific general plan land use classifications that appear in the coverage needs to apply.

5.3 Sources of Research Information

In the past five years MAG has assembled a large body of databases which describe various facets of land use and growth within the region. The key ones are:

- *Existing Land Use Cover (existlu).* The existing land use database is an ARC/INFO polygon database representing land use in 1993. The cover has 13,224 polygons and is coded with four different sets of land use codes (items LU, PCODE, GCODE, MCODE, HPGCODE, and GPLU), some of which appear to be unused. The code selected for use in the SAM is the field LU, and a dictionary for it is shown in Table 5-3.

- *General Plan Cover (gplanlu).* The general plan database is an ARC/INFO polygon coverage that was first built in 1989 through a cooperative effort with local governments in the region and again reviewed and updated in 1992. The

Figure 5-2a
General Plans for Maricopa County

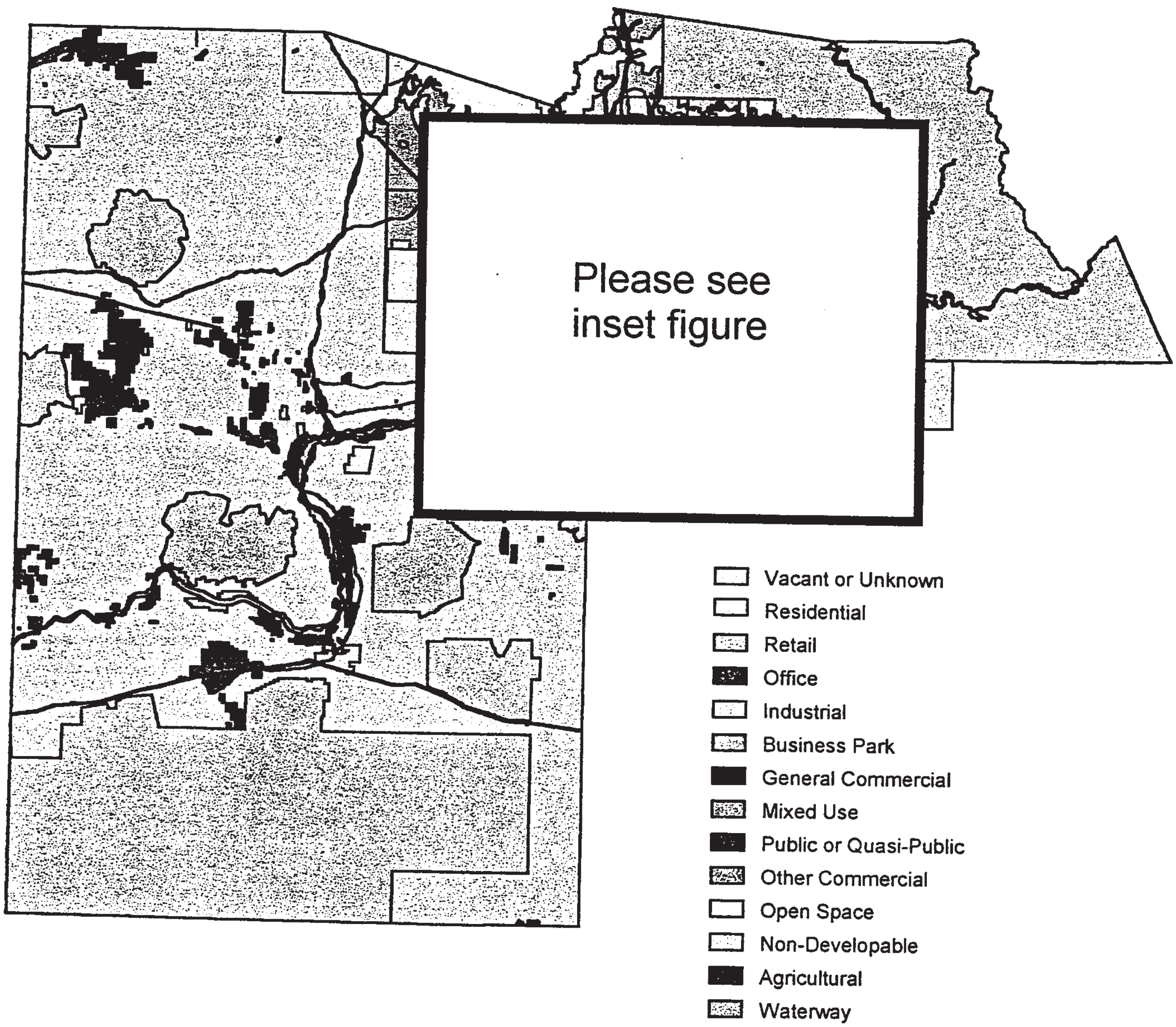


Figure 5-2b

General Plans for Maricopa County (detail)

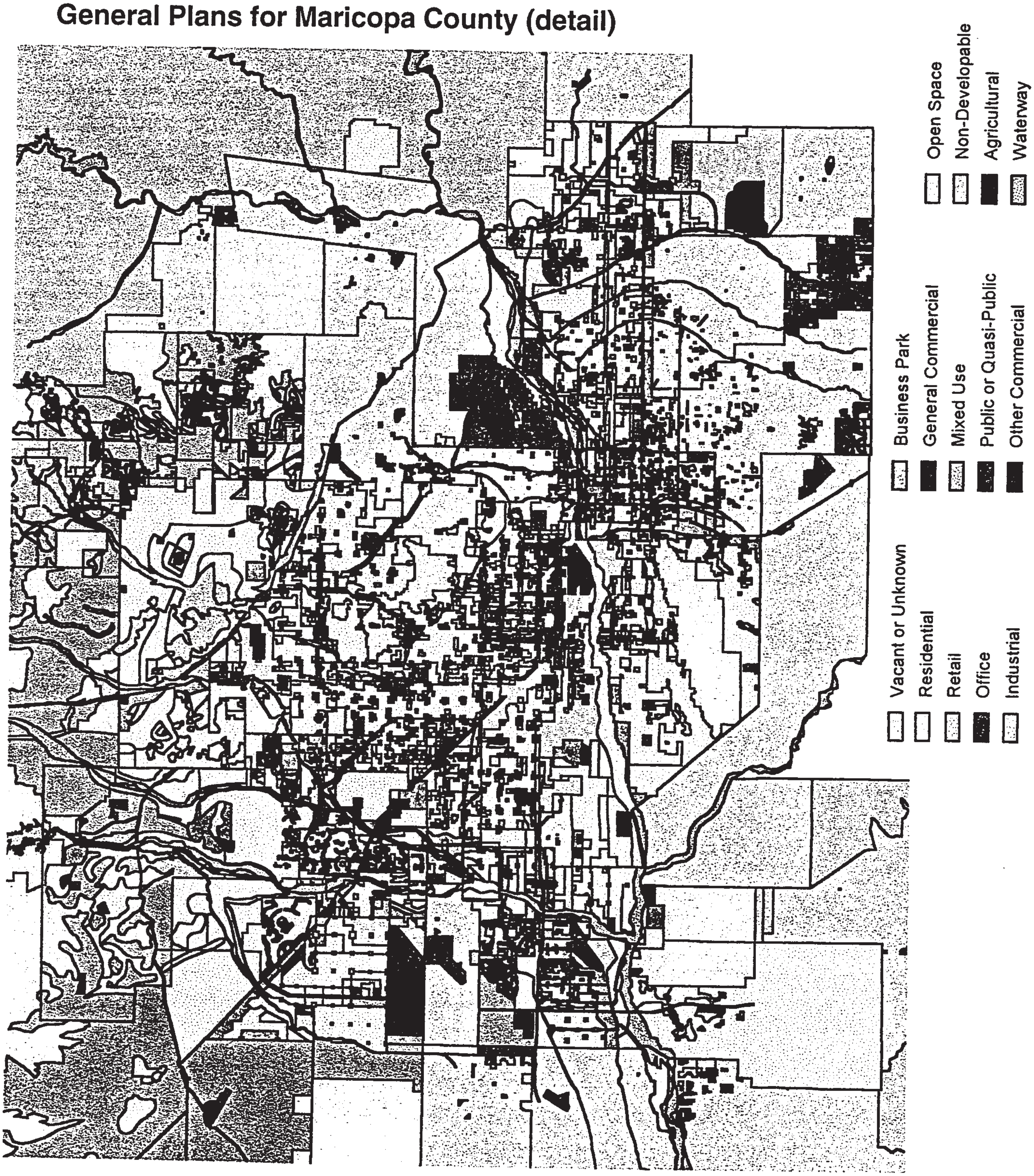


Table 5-2
General Square-Foot-to-Employee Data

Land Use	Employees/1,000 SF	SF per Employee
Neighborhood Retail	1.43	699
Community Retail	1.84	543
Regional Retail	2.26	442
Strip Retail	1.86	538
Retail	2.09	478
Small Offices	3.21	312
Large Offices	2.50	400
Offices	2.86	350
Manufacturing	2.23	448
Warehousing	1.37	730
Public Utilities	4.97	201
Industrial	1.64	610
Schools	1.44	694
Government/Municpal	2.50	400
Public/Quasi Public	1.23	813
Hotel/Motel	2.61	383
Resort	1.96	510
Hotel/Motel/Resort	2.54	394
Hospitals/Clinics	2.80	357
Medical Offices	2.88	347
Medical Services	2.79	358

Table 5-3
Land Use Codes for the Existing Land Use Cover

1 Low Density Residential (SF)	23 Power Stations
2 Medium Density Residential (SF)	24 Railroads
3 High Density Residential (MF)	25 Airports
4 Mobile Homes/RV Parks	
5 Medium Residential Under Development	26 Freeways, Dams
	27 Parks
6 Low Density Commercial	28 Golf Courses
7 Medium Density Commercial	29 Lakes
	30 Rivers
8 Hotel, Resort	31 Vacant Desert
9 Regional Shopping Centers	32 Agriculture - Citrus
	33 Agriculture - Other Crops
10 Commercial Warehouse	34 Agriculture - Stockyards
11 Neighborhood Office Buildings	36 Canals
12 High-Rise Office Buildings	
	40 Nondevelopable - Other
13 Light Industrial	41 Nondevelopable - Forest
14 General Industrial	42 Nondevelopable - Mountain
	43 Nondevelopable - Gunnery Range
15 Unknown	
16 Institutions - Schools	
17 Institutions - Colleges	
18 Institutions - Universities	
19 Institutions - Small Hospitals	
20 Institutions - Large Hospitals	
21 Institutions - Public Facilities	
22 Institutions - Churches	

cover has 6,023 polygons and covers most, but not all, of Maricopa County. There are also a number of land use codes associated with the general plan cover (items GCODE, HPGCODE, HIGH, MED, LOW, and GPLU). The land use codes associated with the general plan item HIGH are quite detailed (they probably reflect a concatenation of all of the different ways that cities describe general plan uses within their own community).

Since it provides the most detailed information (and in fact provides density levels for residential land) and exactly reflects city viewpoints toward growth, it was chosen for use in the SAM. A dictionary for this code is shown in Table 5-4.

- *Parcel Cover 1 (parcels.dbf):* A dBase III file for parcels throughout the region is available and was imported into ARC/INFO as an INFO file. For convenience, this database is partitioned into five separate files; altogether it covers 874,298 parcels. The file was last reviewed and updated by ESG in the

Table 5-4
Land Use Codes for the General Plan Cover

General Plan Land Use	Applicable General Plan Codes
Vacant or Unknown	0, 315, 331, 335
Rural Residential	1, 2, 3, 4, 6, 7, 8, 10, 11, 156, 157, 179, 180, 182, 183, 185
Low Density Residential	14, 17, 18, 19, 20, 21, 22, 23, 24, 26, 27, 28, 44, 158, 159, 189, 190, 202, 301
Medium Density Residential	16, 29, 30, 31, 32, 33, 34, 36, 160, 171, 172, 191, 203, 302, 305
High Density Residential	38, 39, 40, 41, 42, 43, 46, 47, 48, 49, 51, 52, 53, 54, 162, 173, 174, 186, 192, 205, 303, 304
Retail	59, 67, 70, 72, 74, 75, 78, 177, 178, 181, 194, 306, 307
Mixed Use	60, 143, 144, 146, 147, 148, 225
General Commercial	62, 63, 64, 66, 79, 80, 81, 93
Other Commercial	55, 56, 58, 193, 207, 215, 308, 320
Business Park	84, 85, 87, 196
Office	88, 89, 91, 92, 94, 96, 97, 195, 213, 311
Industrial	99, 101, 102, 104, 105, 106, 108, 109, 168, 169, 197, 217, 218, 313, 314
Public or Quasi-public	111, 112, 113, 114, 115, 116, 117, 118, 120, 121, 122, 151, 198, 220, 221, 230, 316, 317, 318, 321, 322, 323, 324, 325, 326
Open Space	126, 129, 130, 131, 132, 133, 134, 135, 136, 137, 140, 223, 224, 327, 328
Non-Developable	125, 138, 199, 222, 340, 341, 342, 343
Waterway	329, 330, 336
Agricultural	332, 333, 334

Socioeconomic Models Enhancement Project in May 1994, and is described fully in their final report. For most parcels, coordinates associated with parcel centroids do not reflect actual parcel location, but instead reflect the centroid of the bookmap where the parcel appears. This database also contains a number of relevant information for the effort, including the following:

- Parcel Number
- Parcel area
- Parcel land use classification code
- A more general MAG land use code
- Number of units (i.e., for residential parcels)
- The date of construction[3]

[3] Actually, several dates associated with construction of structures at the site are included; the YEAREARLIEST field in order to establish the age of the dwelling unit was chosen.

There are questions about the accuracy of the data which appears in the parcel database. It reportedly has not been well maintained over time.

Table 5-5 describes the coding schemes for land use that appear on records in this database.

Table 5-5
Land Use Codes in the Parcel Coverage

Item	Code	Description	Code	Description
MAGLU	01	SF	17	Manufacturing
	02	Townhouse	18	Warehouse
	03	Multifamily	19	Public Utilities
	05	Hotel/Motel	20	Schools
	06	Resort	21	Government
			22	Churches
	07	Neighborhood Retail		
	08	Community Retail	23	Vacant Developable
	09	Regional Retail	24	Vacant NonDevelopable
	10	Strip Commercial	25	Golf Courses
	11	Auto Services/Sales	26	Mining
			27	Under Construction
	12	Small Offices		
	13	Large Offices	28	Other Miscellaneous
	14	Nursing		
	15	Hospitals/Clinics		
	16	Medical Offices		
LCIC	0000-0085	Vacant Land	1000-2900	Commercial Property
	0100-0198	SF Residential	3000-3740	Industrial Property
	0300-0380	MF Residential	4000-4910	Ranch Property
	0400-0600	Hotel/Motel/Resort	5100-5900	Public Utilities
	0700-0770	Condominiums	6100-6900	Natural Resources
	0800-0870	Mobile Homes	8500-8900	Special Use Property
	0900-0950	Miscellaneous and Salvage	9200-9860	General Service Use

- *Parcel Cover 2 (magcen):* A more recent updated version of the parcel file is also available. It reportedly contains more accurate geographic coordinates for the actual parcel centroids themselves, not the bookmap location as is the case with the version described above. There are a total of 907,727 parcels represented in this database, or 33,429 more than in ***parcels.dbf***. It was imported into ARC/INFO and built as a point coverage, called ***magcen***. It, however, includes information only on the parcel number. Therefore, it must be joined with the first version of the parcel coverage described above to obtain other information about the parcels in question.

- *Employment Database (**magemp94.dat**)*: The MAG employment database was also built by ESG as part of the Socioeconomic Model Enhancement project in 1994. It compiles information about *major* employers throughout Maricopa County, that is those employers with 50 or more workers. There are 3,896 such records in the file, covering 490,137 workers, or about 50% of the total estimated employment of 978,667 employees associated with the 1993 trip generation dataset. It also was imported into ARC/INFO and built as the point coverage ***magemp***. Pertinent information which appears in this file includes:

  - Employer identification and location (address and <x,y> coordinate)
  - Number of workers for various years, most recently 1994
  - SIC code

One of the problems associated with these datasets is that there is not a consistent land use coding scheme that applies to all of them.

Table 5-6 compares the depiction of Maricopa County residential land uses as portrayed by the existing land use cover with that portrayed by the parcel coverage (similar land uses together by judgment have been grouped). As is evident from the table, the existing land use cover claims almost three times more land devoted to single family residences than does the parcel coverage.

Table 5-6

Comparison of Residential Land Use Distribution Depicted in the Existing Land Use Cover and the Parcel Cover

Existing Land Use Cover			Parcel Inventory			Percent Ratio
	Land Use	Acres		Land Use	Acres	
	SF Residential					
1	SF (low density)	102,387	1	SF (all)	80,197	
2	SF (medium density)	137,463				
3	Medium (developing)	9,934				
	Total	249,784			80,197	32%
	MF Residential					
3	MF	25,872	3	MF	15,645	60%
	Mobile Home					
4	Mobile Home	11,585	4	Mobile Home	13,567	117%
	Total Residential	287,241			109,409	38%

Table 5-7 provides a similar comparison for non-residential land uses (here the identification of similar land uses in the two classification schemes used for the two coverages is more difficult). In some cases, the two databases roughly agree, but for the most part the two databases appear to offer a substantially different view of land use.

In general, the existing land use cover should be expected to report larger areas devoted to land uses than does the parcel coverage because the existing land use cover reports gross densities — including public rights of way which serve land uses. The parcel coverage, of course, would exclude rights-of-way. Still, the differences between the two portraits of Maricopa County land use cannot be explained entirely by this.

Table 5-7
Comparison of Non-Residential Land Use Distribution Depicted in the Existing Land Use Cover and the Parcel Cover

Existing Land Use Cover			Parcel Inventory			Ratio of Parcel Acres
	Land Use	Acres		Land Use	Acres	Cover Acres
	Commercial/Retail					
6	Commercial (low)	7,858	7	Neighborhood Retail	3,464	
7	Commercial (high)	15,043	8	Community Retail	870	
9	Regional Ctrs	1,211	9	Regional Retail	3,014	
			10	Strip Commercial	9,097	
			11	Auto Services	401	
	Total	24,112			16,845	70%
	Hotels/Motels/Resorts					
8	Hotels/Resorts	1,505	5	Hotels/Motels	1,191	
			6	Resorts	1,140	
	Total	1,505			2,330	155%
	Office					
11	Small Office	2,693	12	Small Office	481	
12	High Rise Office	796	13	Large Office	612	
			16	Medical Office	418	
	Total	3,489			1,511	43%
	Industrial					
10	Commercial Warehouse	1,673	18	Warehouse	7,896	
13	Light Industrial	17,190	17	Manufacturing	7,134	
14	General Industrial	17,347				
	Total	36,210			15,020	41%
	Educational Facilities					
16	Inst Schools	7,057	20	Schools	7,795	
17	Inst Colleges	658				
18	Inst Univers	1,044				
	Total	8,759			7,795	89%

Table 5-7

Comparison of Non-Residential Land Use Distribution Depicted in the Existing Land Use Cover and the Parcel Cover (Continued)

	Existing Land Use Cover			Parcel Inventory		Ratio of Parcel Acres
	Land Use	**Acres**		**Land Use**	**Acres**	**Cover Acres**
	Medical Facilities					
19	Inst: Small Hosp	230	14	Nursing Homes	741	
20	Inst: large Hosp	700	15	Hospitals/Clinics	593	
	Total	930			1,334	143%
	Public Facilities					
21	Inst: Public Fac	5,603	21	Government	112,845	
27	Parks	110,964				
25	Airports	12,320				
22	Inst: Churches	1,789	22	Churches	4,123	
	Total	130,676			116,968	90%
	Utilities					
23	Power Stations	5,202	19	Utilities	14,940	
24	Railroads	2,815				
	Total	8,017			14,940	186%
	Golf Courses					
28	Golf Courses	13,954	25	Golf Courses	8,881	64%
	Agricultural					
32	Agriculture: Citrus	15,895				
33	Agriculture: Other	415,538				
34	Agriculture: Stock	13,535				
	Total	444,786			-0-	n/a
	Vacant					
31	Developable	3,086,087	23	Developable	835,436	
40	Non-Developable	6,379	24	Non-Developable	281,386	
	Total	6,092,466			1,116,822	18%
	Other					
26	Freeways	15,534				
29	Lakes	7,332				
30	Rivers	64,301				
36	Canals	12,589				
	Total	99,756			-0-	n/a

5.4 Housing Analysis

5.4.1 Eligible General Plan Classifications and Absorption Rate

The general plan cover contains designations for residential development, and therefore defines those areas which are eligible for housing growth. Residential designations which appear in the cover are shown in Table 5-8. Documentation for the cover which have been

Table 5-8

Eligible General Plan Codes and Densities for Residential Growth

Code	Type	Dens	Code	Type	Dens	Code	Type	Dens	Code	Type	Dens
1	RR	0.10	29	MDR	4.00	156	RR	0.10	205	HDR	12.50
2	RR	0.27	30	MDR	4.25	157	RR	0.60	301	LDR	1.75
3	RR	0.42	31	MDR	6.00	158	LDR	1.50	302	MDR	5.25
4	RR	0.35	32	MDR	6.50	159	LDR	3.00	303	HDR	16.50
6	RR	0.50	33	MDR	7.00	160	MDR	5.00	304	HDR	16.50
7	RR	0.50	34	MDR	7.50	162	HDR	18.50	305	MDR	5.25
8	RR	0.50	36	MDR	8.00	171	MDR	8.00			
10	RR	0.75	38	HDR	10.00	172	MDR	6.50			
11	RR	0.75	39	HDR	10.00	173	HDR	15.50			
14	LDR	1.00	40	HDR	16.50	174	HDR	15.00			
16	MDR	5.25	41	HDR	10.00	179	RR	0.50			
17	LDR	1.50	42	HDR	10.55	180	RR	0.30			
18	LDR	1.50	43	HDR	10.55	182	RR	0.50			
19	LDR	1.75	44	LDR	4.00	183	RR	0.50			
20	LDR	1.75	46	HDR	12.50	185	RR	0.45			
21	LDR	2.50	47	HDR	13.00	186	HDR	20.00			
22	LDR	3.50	48	HDR	17.50	189	LDR	1.75			
23	LDR	3.55	49	HDR	20.00	190	LDR	1.75			
24	LDR	3.00	51	HDR	7.00	191	MDR	5.25			
26	LDR	3.50	52	HDR	16.00	192	HDR	16.50			
27	LDR	3.50	53	HDR	22.50	202	LDR	3.75			
28	LDR	3.00	54	HDR	25.00	203	MDR	6.25			

RR = Rural Residential
LDR = Low Density Residential
MDR = Medium Density Residential
HDR = High Density Residential

given also expresses density ranges for each of the designations. The midpoint of these ranges in Table 5-8 to establish land absorption rates associated with residential growth have been used. *Note in Table 5-8 that there are no lands in the general plan coverage that designate mobile home communities.*

5.4.2 Distributing Households by Income Class to Dwelling Unit Types

There currently is no information concerning the desirability of various types of dwelling unit types as a function of household income. While the census PUMS dataset provides a loose relationship for the large subregions, substantial differences among neighborhoods and communities in the metropolitan area is expected).

5.4.3 Dwelling Units by Age of Structure

MAG has also requested that the trip generation dataset be augmented to reflect age of dwelling units in TAZs; this means that this information must reside in the SAM.

The *only* source of information for age of dwelling units is the parcel coverage, which records several fields describing construction on parcel sites throughout Maricopa County. These fields include the year of earliest construction at the site, the year most construction took place, and the year of the latest instance of construction at the site. For the purposes of establishing age of dwelling units, the year of *earliest* construction was used. Age classifications were established by MAG, as shown in Table 5-9.

Table 5-9
MAG Classifications for Age from the Parcel File

Age Class	Description	Units	Percent
0	Age missing/unknown	11,911	2.05%
1	Less than 10 years	129,591	22.26%
2	10 to 19 years	167,270	28.73%
3	20 to 29 years	103,950	17.85%
4	30 years or more	169,505	29.11%
	Total	582,227	100.00%

Base Year

Age distributions for each residential polygon in the existing land use cover were constructed from parcel information. A *point-in-polygon* operation was run with ARC/INFO, assigning each residential parcel to the existing land use polygon in which it belonged. (Note that this assumes that the coordinates associated with the parcel centroids are accurate.) Age classifications for each residential parcel were generated based on the earliest construction year at the site (and the base year 1995). Distributions were then run (weighted by the number of units associated with each residential parcel) and aggregated for each land use polygon.

In fact, since the data on parcels spanned two different databases (one with attribute information about parcels, the other with location), the two parcel coverages were first *joined* before the analysis was carried out.

Future users of this dataset should be aware that the quality of information available from the parcel coverage is somewhat suspect and therefore the analysis should probably be repeated when better information becomes available. These are the problems which were noted during the analysis:

- The number of units associated with single family parcels was almost always blank —one unit for these parcels is assumed.

- The field recording the number of units for multiple family parcels was 6-characters wide, but data was left justified and 0-filled. This means that there is no way to distinguish whether the value "600000" reflects 6 units, 60 units, 600 units, or whatever. In these cases, a single digit value is assumed.

The result of this analysis is an age distribution of dwelling unit age applicable to each existing land use polygon. When gridded, each residential grid cell will inherit the age distribution as an attribute.

Future Residential Age

During forecasts, future residential age associated with grid cells can be handled by simply aging the dwelling unit age distribution associated with the cell, as follows:

- Ages will be assumed to be uniformly distributed throughout the decade classification scheme. If it has been five years since the previous forecast target year, then 50% of the households in any given classification can be assumed to age into the next classification.

5.5 Employment Analysis

5.5.1 Density Levels and Absorption Rates

Numerous attempts were made to establish typical density levels for different land use types which appear in the general plan coverage. While MAG datasets provide substantial information about land uses and employment, no one dataset provides all of the information that is needed. For example:

- ***magemp94***: The survey of major employers (50 employees or more) that was conducted in 1993-4 by the Economic Strategies Group (ESG) resulted in the creation of an inventory database covering about 50% of the total employment in the region. Problems are, however:
    - Businesses are classified econometrically according to SIC, not by land use. Therefore, there is no way to relate whether an mining concern, for example, is actually associated with office land use (because it is the corporate headquarters).
    - The database contains information on total employment, but there is no information on parcel size acreage. Therefore, there is no way to relate employment with underlying land absorption rates. The database *does* contain information on building square footage, but again without some additional data on floor-area ratios or parcel size, the database cannot support the computation of land absorption rates.
- ***parcels***: The parcel inventory does contain information about land use classification (in two ways, the so-called MAGLU number and the Maricopa

County land use classification index LCIC). Unfortunately, there is no information on employment, and as a consequence the parcel inventory does not yield information on employment densities.

A number of attempts were made to derive land absorption rates for employment. For example:

- An attempt was made to match the employer database with the parcel inventory, which *does* report parcel size (based on a *nearest neighbor* GIS operation). This attempt failed because employer locations typically matched the wrong parcels (as determined by identifying the LCIC and MAGLU codes from the parcel). The reasons for these mismatches are unclear, but may include:

  - Parcel locations may be centroids, whereas employer locations were address-matched and therefore are associated with an interpolated point along the street to which the address points.

  - Either, or both, databases may be inaccurate.

- An attempt was made to overlay employer locations on the existing land use cover in order to estimate employment associated with the land use polygons that appear as existing land use. (Granted, the fact that magemp94 only records about 50% of the employment in the county would be a significant source of error).

This last effort generated the results shown in Table 5-10. One issue that the analysis raises is that there are a substantial number of major employers who appear to reside in inappropriate land uses, such as residential uses.[4] While in general there certainly would be employment associated with residential land uses (work-at-home), it is less likely that major employers with 50 or more workers would be found there. The reasons why this occur might include:

- The existing land use cover is generalized, that is it does not account for occasional uses contrary to the overall prevailing designation of land.

- Possibly major employer addresses were address-matched without offsets, meaning that the actual coordinate falls on street centerlines and on boundaries between land use polygons. In these cases, ARC/INFO will assign the point to one of the adjacent polygons sharing the boundary arbitrarily.

In any event, these anomalies should be investigated further in the future. Heavy reliance on these database as the foundation for land use modeling at MAG will require more accurate information.

---

[4] A GIS analysis overlaying the parcel file on existing land use was performed. This also generated the same conflicts between land use reported in the parcel coverage with that reported in the existing land use cover.

Table 5-10

Overlay of MAG Major Employer Inventory (*magemp94*) and the Existing Land Use Cover

Cde	Land Use	Polygons	Acres	Avg Size	%	Emp	Dens
1	SF (Low density)	1,073	102,387	95.42	1.73%	6,037	0.06
2	SF (Medium)	904	137,463	152.06	2.32%	28,445	0.21
3	MF (High)	1,100	25,872	23.52	0.44%	13,648	0.53
4	Mobile Homes/RVPark	272	11,585	42.59	0.20%	930	0.08
5	Medium Res/Under Dev	335	9,934	29.65	0.17%	2,334	0.23
6	Low Density Comml	748	7,858	10.51	0.13%	15,952	2.03
7	Medium Dens Comml	587	15,043	25.63	0.25%	73,756	4.90
8	Hotel/Resort	95	1,505	15.85	0.03%	11,571	7.69
9	Regional Shopping Ctr	25	1,211	48.45	0.02%	3,551	2.93
10	Comml Warehouse	54	1,673	30.97	0.03%	5,961	3.56
11	Small Office	297	2,693	9.07	0.05%	29,474	10.94
12	High Rise Office	47	796	16.93	0.01%	33,258	41.79
13	Light Industrial	431	17,190	39.88	0.29%	60,010	3.49
14	General Industrial	186	17,347	93.26	0.29%	59,136	3.41
15	Unknown	416	2,302	5.53	0.04%	616	0.27
16	Inst: Schools	436	7,057	16.19	0.12%	21,564	3.06
17	Inst: Colleges	14	658	46.98	0.01%	1,990	3.03
18	Inst: Universities	7	1,044	149.13	0.02%	901	0.86
19	Inst: Small Hosp	21	230	10.97	0.00%	3,277	14.22
20	Inst: Large Hosp	21	700	33.34	0.01%	22,510	32.15
21	Inst: Public	244	5,603	22.96	0.09%	9,477	1.69
22	Inst: Churches	323	1,789	5.54	0.03%	982	0.55
23	Power Stations	146	5,202	35.63	0.09%	492	0.09
24	Railroads	44	2,815	63.99	0.05%	1,373	0.49
25	Airports	23	12,320	535.65	0.21%	13,222	1.07
26	Freeways/Dams	82	15,534	189.44	0.26%	3,715	0.24
27	Parks	581	110,964	190.99	1.87%	4,204	0.04
28	Golf Courses	375	13,954	37.21	0.24%	5,401	0.39
29	Lakes	274	7,332	26.76	0.12%	95	0.01
30	Rivers	186	64,301	345.70	1.08%	383	0.01
31	Vacant Desert	2,037	3,086,087	1515.02	52.00%	44,474	0.01
32	Agr: Citrus	181	15,895	87.82	0.27%	497	0.03
33	Agr: Other	649	415,538	640.27	7.00%	9,732	0.02
34	Agr: Stockyards	848	13,353	15.75	0.22%	810	0.06
36	Canals	143	12,589	88.03	0.21%	1,964	0.16
40	Nondevelopable	5	6,379	1275.74	0.11%		0.00
41	Nondevelopable	3	646,507	215502.31	10.89%		0.00
42	Nondevelopable	9	315,996	35110.67	5.32%		0.00
43	Nondevelopable	1	818,507	818506.89	13.79%		0.00
	Total	13,223	5,935,212	448.861			491,742

5.5.2 General Plan Non-Residential Land Absorption Rates

Despite these data deficiencies, this methodology offered the only opportunity for establishing gross land absorption rates for non-residential general plan land use classifications. The method is as followed:

- First, one of the DRAM/EMPAL employment types to each of the existing land use classifications reported in the existing land use cover was assigned. For example, by definition that the office land use was entirely populated by nothing but office employees was assumed. These assignments for each land use are shown in Table 5-11.

- Once accomplished, the types of employment was totaled (office, retail, industrial, public, and other) that appear in the major employer database. (Recall that the SIC codes included in this database do not relate to the land use definitions which appear in the land use cover). For each of the five categories of employment, adjustment factors to account for the fact that only 50% of regional employment appears in the major employer database overall was obtained. These factors are shown in Table 5-12.

- Then adjusted each employment type upwards accordingly. This assumes that there is not any geographic bias associated with omitting small employers in the database, somewhat a leap of faith.

- The resulting estimates of employment densities for existing land uses, then, is shown in Table 5-11.

The final step involves assigning these densities to general plan land use classifications. This was done by judgment, relating each land use classification in the general plan coverage to representative classifications that seemed suitable in the existing land use cover. These representative land use classifications were used to generate a prevailing land absorption rate (employment density) that are shown in Table 5-13. Until better information becomes available, these will be used in the SAM.

5.6 Locational Preference Criteria

The SAM will also require some way to rate individual lands in terms of their priority for development. This mechanism could be seen to become increasingly sophisticated as the SAM is used and developed in the future, based partly on the performance of the model and partly on additional research that might be undertaken to study this further. For now, what appears to be a reasonable mechanism would involve the following ratings for both residential and non-residential land uses.

1. Proximity to major highways, that is, those appearing in the network associated with the forecast year.

2. General proximity to developed land, that is, the notion that cities tend to grow outward.

3. Immediate proximity to developed land, that is lands immediately adjacent to development tend to be developed earlier than lands that are more remote.

Table 5-11

Employment Density Estimates for Existing Land Use Codes

Cde	Land Use	Acres	Estimated Employment	Average Density	Employment Type
1	SF (Low density)	102,387	6,547	0.064	Other
2	SF (Medium)	137,463	30,848	0.224	Other
3	MF (High)	25,872	14,801	0.572	Other
4	Mobile Homes/RVPark	11,585	1,009	0.087	Other
5	Medium Res/Under Dev	9,934	2,531	0.255	Other
6	Low Density Comml	7,858	40,595	5.166	Retail
7	Medium Dens Comml	15,043	187,696	12.477	Retail
8	Hotel/Resort	1,505	12,549	8.336	Other
9	Regional Shopping Ctr	1,211	9,037	7.461	Retail
10	Comml Warehouse	1,673	10,820	6.469	Industrial
11	Small Office	2,693	112,151	41.647	Office
12	High Rise Office	796	126,549	159.005	Office
13	Light Industrial	17,190	108,928	6.337	Industrial
14	General Industrial	17,347	107,341	6.188	Industrial
15	Unknown	2,302	668	0.290	Other
16	Inst: Schools	7,057	77,671	11.006	Public
17	Inst: Colleges	658	2,158	3.281	Other
18	Inst: Universities	1,044	3,258	3.109	Public
19	Inst: Small Hosp	230	3,554	15.425	Other
20	Inst: Large Hosp	700	24,412	34.867	Other
21	Inst: Public	5,603	34,135	6.093	Public
22	Inst: Churches	1,789	1,065	0.595	Other
23	Power Stations	5,202	893	0.172	Industrial
24	Railroads	2,815	2,492	0.885	Industrial
25	Airports	12,320	24,000	1.948	Industrial
26	Freeways/Dams	15,534	0		0
27	Parks	110,964	15,142	0.136	Public
28	Golf Courses	13,954	5,857	0.420	Other
29	Lakes	7,332	0		0
30	Rivers	64,301	0		0
31	Vacant Desert	3,086,087	0		0
32	Agr: Citrus	15,895	539	0.034	Other
33	Agr: Other	415,538	10,554	0.025	Other
34	Agr: Stockyards	13,353	878	0.066	Other
36	Canals	12,589	0		0
	Total	5,935,212			978,667

Table 5-12
Factors to Create Estimated Employment by Type

Employment Type	From Maj. Emp.	1994 Estimated	Factor
Other	108,780	117,971	1.08
Public	36,146	130,194	3.60
Retail	93,259	237,328	2.54
Office	62,732	238,700	3.81
Industrial	140,194	254,474	1.82
Total	491,742	978,667	

Table 5-13
Derived General Plan Employment Densities

General Plan Class	Representative Existing LU Codes	Rep Acres	Rep Employment	Density
Rural Residential	1	102,387	6,547	0.064
Low Dens Residential	1	102,387	6,547	0.064
Med Dens Residential	2,4,5	158,982	34,388	0.216
High Dens Residential	3	25,872	14,801	0.572
Retail	6,7,9	24,112	237,328	9.843
Mixed use	6,7,9,8,11	28,311	362,027	12.788
General Commercial	6,7,9	24,112	237,328	9.843
Other Commercial	6,7,9	24,112	237,328	9.843
Business Park	11,13	19,882	221,078	11.119
Office	11,12	3,489	238,700	68.419
Other Office	11	2,693	112,151	41.647
High Rise Office	12	796	126,549	159.005
Industrial	10,13,14	36,209	227,089	6.272
Public	16,17,18,19,20	9,689	111,040	11.460
Agriculture	32,33,34	444,786	11,972	0.027

5.7 Recommendations for the Future

The analysis was undertaken in this task raises a number of issues which should be addressed, especially with the growing reliance on these datasets for modeling (including the emergence of subarea focused transportation models in the region):

- Land uses and employment types need to be more coherently defined. A common or compatible coding system is required.
- Apparent anomalies between the various datasets need to be investigated and resolved.
- Future employment surveys should include elements to determine (1) parcel sizes of employers at the sites they occupy and (2) a determination of the land use at the site (not the SIC code of the business).

6
Update Special Population Group Methodologies

The MAG trip generation model is driven by a number of demographic variables which describe land use (actually, socio-economic) patterns in Maricopa County. The variables include various measures of population, households, income, and employment. In addition, the trip generation model also requires information about three population groups: **seasonal population, transient population,** and **group quarters population.** These are known collectively as the *special population groups* as distinguished from permanent regular residents of the region which are forecast separately in DRAM/EMPAL. That is, forecasts for special population groups do not come from DRAM/EMPAL.

Methods for forecasting special population groups is the subject of this section.

6.1 Task Objectives

In 1993 the Economic Strategies Group conducted a considerable body of research into methodologies for generating forecasts of special population groups. The objective in this task was to review those methodologies and recommend any apparent revisions needed.

However, one of the overall objectives of the project concerned the need to develop a Subarea Allocation Model (SAM) that would enable MAG planners to disaggregate DRAM/EMPAL forecast data (which does not address special population groups) which is generated on a RAZ basis, to TAZs. Later, in view of MAG's desire to be able to generate forecast data for any geography, the goal of the SAM was revised to enable MAG to disaggregate all forecast data to a Minimal Analysis Zone (MAZ) system, thereby giving MAG the capability to report data on any geography through simple aggregation.

Consequently, it became apparent that the SAM capability should not only address socio-economic forecast data generated by DRAM/EMPAL, but also should address special population groups. Therefore, an important objective of this task was to build an ARC/INFO capability for disaggregating special population group forecast control totals for all of Maricopa County to the MAZ level.

6.2 Background on Existing Methods

The special population groups are defined to include:

- *Seasonal Population:* Seasonal populations are those nonresidents that reside in the region at certain times of the year for periods greater than two weeks.

- *Transient Population:* Transient populations include nonresidents that reside in the region for periods less than two weeks (it would therefore describe business travelers and tourists).

- *Group Quarters Population:* Group quarters include that portion of the resident population that reside in non-household quarters, such as institutions, congregate care facilities, dormitories.

More elaborate definitions are available in the **Socioeconomic Model Enhancement Final Report** by ESG, dated May 1994. Also, a full discussion of the forecast methodologies which were developed by ESG related to these three special population groups is also provided in this report.

A brief review of the methods will now be given. The intent is not to repeat the extensive discussion about the derivation and justification for the methods given by ESG that appear in their final report. Instead, only describe the methodology developed so it can be adapted to use in the SAM.

Note that an important part of the methodology not only addresses the need for forecasting *socio-economic* descriptions of the county (e.g., seasonal and transient populations), but also addresses the need to translate those descriptions into *land use* forecasts (e.g., motels, nursing homes).

6.2.1 Seasonal Population

Research conducted by ESG defined seasonal populations to include parts of populations associated with mobile home parks housing, recreational vehicle/travel trailer parks, and general dwelling units (e.g., retirement housing). While ESG referred to the last class of seasonal populations as *other*, they in fact seem to treat them to be specifically residents of formal *retirement communities*. Therefore, use that terminology.

Seasonal visitors, then, were categorized by the three types of land uses, mobile homes, RV parks, and retirement housing. Then the countywide forecasts were disaggregated to the RAZ level, also taking into consideration *known* projects.

County-Wide Forecast Methodology

The ESG methodology generally is as follows:

- Total seasonal populations residing in dwelling units (i.e., not including those in RV parks) in the future will be as measured in the census, 5.905% of total Maricopa County housing, with an average of 2.0 persons per dwelling unit. These projections are shown below in Table 6-1.

Table 6-1
Seasonal Population Forecasts in Dwelling Units
(Taken from Table IV-15 in the ESG Final Report)

	Forecast		
Year	Maricopa County Housing	Percentage Seasonal	Total Seasonal Housing
1990	952,041	5.905%	56,220
1995	1,076,699	5.905%	63,581
2000	1,222,020	5.905%	72,163
2015	1,687,439	5.905%	99,647
2040	2,674,486	5.905%	157,934

- The portion of seasonal housing residing specifically in *mobile homes* can be computed from the ASU Winter Visitor survey, which identified state of origin for all seasonal visitors staying in mobile homes. Visitation rates from each state reported in the survey were computed based on population age 65 and above in those states. Forecasts of populations 65 and above in those states provide the basis for estimating total seasonal visitors from them in the future, assuming that the same rates apply. An occupancy of 2.0 persons per unit was assumed. Total occupied mobile homes (including regular household population) was also taken from the survey, which determined that 45% of all occupied mobile homes were residences for seasonal visitors and maximum overall mobile home occupancy was 87%.

  Therefore, seasonal populations residing in mobile homes were forecast as indicated in Table 6-2.

- Total *seasonal visitation* in retirement communities was forecast simply based on the difference between overall seasonal visitation in fixed dwelling units reported in Table 6-1 and those in mobile homes (Table 6-2). These are shown in Table 6-3.

Table 6-2
Mobile Home Forecasts for Seasonal Population
(from ESG Report Chapter IV; some variances from report shown here)

Year	Seasonal Visitation in Mobile Homes [(1)]	Seasonal Mobile Home Units [(2)]	Total Occupied Mobile Home Units [(3)]	Total Inventory of Mobile Home Units [(4)]
1992	44,185	22,093	49,094	56,430
1995	45,364	22,682	50,404	57,936
2000	46,714	23,357	51,904	59,660
2015	61,132	30,566	67,924	78,074
2040	97,207	48,604	108,008	124,147

Notes:
(1) Based on per capita visitation rates from other states and Canada for forecast populations aged 65 and above in those states.
(2) Based on an average of 2.0 persons/unit.
(3) Based on 45% of mobile homes occupied by seasonal visitors.
(4) Based on maximum occupancy rate for mobile homes of 87%.

Table 6-3
Seasonal Population in Retirement Communities (Other Seasonal Population)
(from ESG Report Chapter IV; some variances from report shown here)

Year	Total Seasonal Housing [(1)]	In Mobile Homes [(2)]	In Retirement Communities [(3)]	Retirement Community Population [(4)]
1990	56,218	22,093	34,126	68,252
1995	63,579	22,682	40,897	81,794
2000	72,160	23,357	48,803	97,606
2015	99,643	30,566	69,077	138,154
2040	157,928	48,604	109,325	218,650

Notes:
(1) From Table 6-1
(2) From Table 6-2
(3) (1) - (2)
(4) Assuming 2.0 persons/household

- The portion of seasonal housing residing in *RV parks* followed a rationale similar to that used for mobile homes, using measures specifically for RV parks obtained during the ASU Winter Visitor survey. Forecast visitation was also predicated on population aged 65 and above from origin states throughout the country. Per capita visitation from those states was assumed to be as measured in the survey. Seasonal population associated with RV parks is shown in Table 6-4.

  Total occupied RV spaces assumed 2.0 persons/space, 82% occupancy by out-of-state seasonal visitors, and a maximum occupancy level of 90%.

Table 6-4
RV Park Forecasts for Seasonal Population
(from ESG Report Chapter IV; some variances from report shown here)

Year	Seasonal Visitation in RV Parks (1)	Seasonal RV Park Units (2)	Total Occupied RV Park Spaces (3)	Total Inventory of RV Park Units (4)
1992	50,072	25,036	30,532	33,924
1995	51,407	25,704	31,346	34,829
2000	52,937	26,469	32,279	35,865
2015	69,275	34,638	42,241	46,934
2040	110,154	55,077	67,167	74,630

Notes:
(1) Based on per capita visitation rates from other states and Canada for forecast populations aged 65 and above in those states.
(2) Based on an average of 2.0 persons/unit.
(3) Based on 82% of RV spaces occupied by seasonal visitors.
(4) Based on maximum occupancy rate for RV Parks of 90%.

Consequently, note that the ESG methodology forecasts seasonal visitation based on census and survey statistics and converts those forecasts into land use and development associated with it. These land use categories are used to disaggregate the countywide forecasts to individual RAZs.

Total seasonal populations encompassing these three classes, then, are shown in Table 6-5.

Disaggregation of Seasonal Population

Countywide control totals for forecast seasonal population (by land use type) are disaggregated to individual RAZs based first on known projects (retirement community plans and mobile home/RV projects). The balance of growth associated with the countywide forecast is then prorated among RAZs (although the report does not discuss on what basis). ESG's efforts to develop a statistical analysis for locating growth was not successful.

Table 6-5
Total Seasonal Populations

Year	Mobile Homes [1]	RV Parks [2]	Retirement Communities [3]	Total [4]
1990	n/a	n/a	68,252	n/a
1995	45,364	51,407	81,794	178,565
2000	46,714	52,937	97,606	197,257
2015	61,132	69,275	138,154	268,561
2040	97,207	110,154	218,650	426,011

Notes:
[1] From Table 6-1
[2] From Table 6-3
[3] From Table 6-2
[4] = (1) + (2) + (3)

Data on retirement communities is shown in Table 6-6. Data on known mobile home and RV park projects is shown in Table 6-7. Note that holding capacity of existing retirement communities listed in Table 6-6 will permit growth on the order of 17,537 units, which will carry projected demand in Maricopa County through the year 2015. Afterwards, new retirement communities will have to arise to accommodate demand.

Table 6-6
Known Retirement Community Projects
(from ESG Final Report, Table VII-1)

Community	RAZ	Year Founded	Total Acres	Residential Acres	Dwelling Unites Planned	Dwelling Units Built	Buildout Population
Sun City	237	1960	8,900	n/a	27,353	Built Out	38,126
Sun City West	221	1978	6,575	3,600	15,500	11,537	21,700
Sun Lakes	325/326	1973	3,322	2,441	12,800	4,143	19,200
Dreamland Villa	299	1959	640	n/a	5,000	Built Out	9,500
Westbrook Village	215	1983	1,326	967	4,000	2,533	5,600
Leisure World	299	1973	1,120	495	2,564	Built Out	4,500
Sunland Village	299	1974	570	570	2,549	Built Out	3,608
Sunland Village East	321	1985	582	446	2,491	1,192	3,487
Fountain of the Sun	300	1972	582	454	2,309	2,190	3,233
Sunbird Golf Resort	328	1987	652	320	1,717	750	2,404
Youngtown	236	1954	717	493	1,670	Built Out	2,542
Sun Village	232	1988	335	231	1,356	692	2,500
Rio Verde	231	1973	709	544	1,051	650	2,812
Total			26,030		80,360	62,823	119,212

Table 6-7
Known Mobile Home and RV Park Projects
(from ESG Final Report, Page VII-13)

Category	RAZ	TAZ	Units	City	Project
Mobile Home Parks	331	n/a	30	Gila Bend	North of I-8/South of Highway 85
	201	n/a	600	Wickenberg	Adjacent to American Inn, East of Highway 60
RV Parks	232	145	200	Surprise	Expansion of Happy Trails at 17200 W Bell Rd.

6.2.2 Transient Population

Transient populations are those presumed to reside in hotels and motels. The ESG methodology for generating estimates of transient population is as follows:

- Transient population was subclassified into four component groups, which included (1) leisure travelers from domestic and foreign lands, (2) group travellers, (3) business visitors, and (4) other visitors. Each component of *average daily visitation* was forecast separately.

- Hotel and motel room inventory required to house visitors was then estimated, based on an assumed persons/room rate associated with each component of visitation. Total hotel/motel room inventory was then estimated based on a predicted gradual increase in overall occupancy rate throughout the region of 59.71% in 1990 to 70.00% in the year 2040.

- In addition, peak visitation was projected to be 36.34% above average visitation, across the board for *all* visitor classes, for *all* years.

- Visitors were then assigned to individual RAZs and TAZs based on (1) an inventory of hotel/motel space and (2) known hotel/motel projects in the region. For the out years, for which there are no known hotel/motel projects, forecast growth in the visitation market was pro-rated, although the details of this pro-rated method were not described.

County-Wide Forecast Methodology

Methods for forecasting countywide control totals for total visitation, based on each of the four components of visitation, were as follows:

- *Leisure Visitation:* A method for forecasting average daily leisure visitation was developed in consideration of existing state and country of origin available from local studies of tourism. Forecast population growth in those states and countries was then used to develop forecasts to Maricopa County.

- *Group Visitation:* Daily visitation for group visitors was assumed to from domestic sources only, however forecast in employment growth in states was used as a basis for determining state of origin (instead of population as was the case for leisure visitors). No information concerning state of origin for group visitors was available, so the breakdown by state for leisure visitors was used.

- *Business Visitors:* Business visitors were forecast based on 0.977% of total Maricopa County employment.

- *Other Visitors:* Other visitors were forecast based on 4.93% of the total of the other three visitor categories.

Table 6-8 summarizes the forecasts of visitation to the County developed by ESG.

Table 6-8

Visitation Forecasts to Maricopa County

(from various tables, ESG Final Report)

	Leisure Visitors							
Year	Domestic (1)	Foreign (1)	Total (1)	Group Visitors (2)	Regional Employmt	Business Visitors (3)	Other Visitors (4)	Total Visitors (5)
1990	11,197	1,527	12,724	6,373	975,037	9,535	1,412	30,043
1993	11,499	1,603	13,102	6,627	1,061,144	10,377	1,484	31,590
1995	11,701	1,653	13,354	6,796	1,100,082	10,758	1,524	32,431
2000	12,118	1,777	13,895	7,179	1,288,659	12,602	1,660	35,336
2015	13,117	2,123	15,240	7,512	1,691,143	16,538	1,937	41,227
2040	13,784	2,565	16,349	7,243	2,370,515	23,181	2,306	49,079

Notes:
(1) Leisure visitors forecast independently.
(2) Group visitors forecast independently.
(3) Business visitors based on 0.977% of regional employment.
(4) Other visitors based on 4.93% of other three categories of visitation.
(5) = (1) + (2) + (3) + (4)

These forecasts were then converted into forecasts of hotel/motel/resort rooms as shown in Table 6-9.

Table 6-9

Forecast Hotel/Motel Rooms in Maricopa County

(from ESG Final Report; page VII-13)

Year	Leisure [1]	Occupied Rooms Group [2]	Business [3]	Other [4]	Total Rooms [5]	Occupancy Rate [6]	Inventory Rooms [7]
1990	6,362	4,721	9,535	941	21,559	59.17%	36,435
1993	6,551	4,909	10,377	989	22,826	59.82%	38,158
1995	6,677	5,034	10,758	1,016	23,485	60.25%	38,977
2000	6,948	5,318	12,602	1,107	25,974	61.34%	42,347
2015	7,620	5,564	16,538	1,291	31,013	64.59%	48,020
2040	8,175	5,365	23,181	1,537	38,258	70.00%	54,655

Notes:
(1) At 2.00 persons per room
(2) At 1.35 persons per room
(3) At 1.00 persons per room
(4) At 1.50 persons per room
(5) = (1) + (2) + (3) + (4)
(6) Overall regional hotel/motel occupancy assumed to gradually increase from 59.17% in 1990 to 70.00% in the year 2040. It appears that ESG used a straight-line increase for intermediate years.
(7) (5) / (6)

Disaggregation of Transient Population

Transient populations related to hotels and motels were disaggregated to the RAZ geography first by considering known projects, which are shown in Table 6-10. Then, a *locational preference model* was built and used to distribute transient populations. The equations in this model were:

Leisure (r^2 = 0.908)

$$P_{leisure}=0.3838P_{transient(t-1)}+0.0197Index_{retail}-0.0754Index_{recreational}+19.20$$

Group (r^2 = 0.68)

$$P_{group}=0.3122P_{transient(t-1)}+0.5824Dens_{pop}-1.0691Dens_{emp}+5.74$$

Business (r^2 = 0.895)

$$P_{business}=0.2733P_{transient(t-1)}+0.2639Dens_{emp}+0.0458Index_{airport}+6.52$$

Table 6-10

Known Hotel/Motel/Resort Projects in Maricopa County

(Source: ESG Final Report)

RAZ	TAZ	Units	City	Project
324	1161	200	Avondale	Near Phoenix International Raceway, depending on delivery of infrastructure and Gila River Bridge crossing.
277	n/a	80	Buckeye	at Junction of Highway 85 and Oglesby Road; needs water supply.
207	5	200	Cave Creek	proposed resort, golf course already open.
315	n/a	200	Chandler	City seeking proposals for a downtown or I-10 corridor hotel. Additional demand from Casino gambling or industrial expansion near Queen Creek Rd/I-10
250	265,334	50	Fountain Hills	Sites have been designated, none built; Saguaro Drive/Shea Blvd most likely initial site
331	1380	60	Gila Bend	Refurbished motel to reopen in October, 1993 at 1046 Pima Street
318	1153	200	Gilbert	SW corner of Warner/Gilbert
311	1074	400	Gilbert	McQueen/Guadalupe
258	371	100	Glendale	expansion of Sage Hotel at 5949 NW Grande
222	156	200	Glendale	at Bell/79th near Arrowhead Mall
309	1016	256	Mesa	in connection with existing Hilton at Alma School/Superstition Freeway
263	388	500	Paradise Valley	proposed resort at Indian Bend/Scottsdale Rd
262	420	225	Paradise Valley	Two proposed expansions to Mountain Shadows at 56th/Lincoln and Camelback Inn at 54th/Lincoln.
262	387	300	Paradise Valley	at Scottsdale/Hummingbird
275	756	500	Phoenix	Downtown at 2nd/3rd north of Jefferson adjacent to City Hall
273	598	100	Tolleson	Westcor high rise north of I-10/99th
273	674	400	Tolleson	Another proposal

Table 6-10

Group Quarters Forecasts for Maricopa County

(from Table IV-1 ESG Final Report)

Year	Population	Age 18-22 (DES)	Age 65+ (DES)	Miltry [1]	Jails [2]	College Dorms [3]	Nursing Homes [4]	Other [5]	Total [6]
1990	2,122,101	n/a	n/a	1,316	8,472	5,256	8,659	9,607	33,310
1993	2,285,199	n/a	12.358%	1,616	9,123	4,242	9,206	10,345	34,232
1995	2,399,600	6.590%	12.156%	1,002	9,580	4,868	9,510	10,863	35,823
2000	2,715,097	6.554%	11.320%	1,002	10,839	5,487	10,020	12,292	39,640
2015	3,724,105	7.422%	11.213%	1,002	14,868	8,509	13,613	16,859	54,851
2040	5,807,906	7.104%	14.912%	1,002	23,187	12,702	28,235	26,293	91,419

Notes:
(1) Military assumes no growth
(2) Jails/prisons assumes 0.399% of total population
(3) College dorms assumes 3.08% of age 18-22 cohort
(4) Nursing homes assumes 3.26% of age 65+ cohort
(5) Other assumes 0.453% of total population
(6) = (1) + (2) + (3) + (4) + (5)

Readers will have to refer to the actual details of the spreadsheets implementing the locational preference models to learn more about their implementation, the units, etc. Note the following:

1. The models are based on *locational indices*, measuring the (airline) distances of RAZs from various attractor sites.
2. The models are *recursive*, in other words the results depend on the distribution of transient population during the last iteration (that is, the previous forecast year).
3. The reported *correlations* (r^2) are high, although this is for highly aggregate RAZ data.
4. The models were calibrated, and only work for, RAZs.

6.2.3 Group Quarters Population

Group quarters populations encompass five components: (1) military base barracks housing, (2) jails and prisons, (3) college and university dormitories, (4) nursing homes, and (5) others (such as mental and juvenile institutions and the homeless).

The ESG methodology first forecasts county-wide control totals for each category of group quarters population. It then disaggregates these control totals to RAZs (and possibly TAZs too, although this is not mentioned).

County-Wide Forecast Methodology

These are the methods for generating group quarters population forecasts for Maricopa County:

- Nursing Homes: 3.26% of population aged 65 and over
- Universities: 3.0785% of DES projections of County population, ages 18 to 22
- Jails and Prisons: 0.399% of total population
- Military Housing: No growth at Luke AFB (closing of Williams AFB reflected)
- Other: 0.45% of total population

Population forecasts for future years can be derived from DRAM/EMPAL. Information on age cohorts comes from the Department of Economic Security (DES). Table 6-11 shows how control total forecasts for Maricopa County are generated for group quarters population.

Disaggregation of Group Quarters Populations

ESG prepared inventories of group quarter facilities in Maricopa County, one of which (nursing homes) resulted in the creation of a database. ESG also conducted a review of known or anticipated developments in the region for each land use category. These are shown in Table 6-12.

ESG reported that regression analysis could not be used to predict the location of group quarters populations in the future because the sample was too small.[1] Put another way, group quarters populations involve several discrete instances associated with the institutions involved, and therefore does not lend itself very well to a regression type analysis.

Therefore, forecasts of group quarters were assigned first to known projects, and then were distributed to RAZs on a *"pro-rata basis."* It is unclear from the report exactly what pro-rata means (that is, on the basis of population at large, then-existing group quarters populations, or what).

[1] ESG reported some success with nursing homes, declaring that proximity to hospitals and to population density tested to be adequate predictors.

Table 6-11
Known Group Quarters Projects
(from ESG Final Report, Chapter VII)

Group Quarters Type	RAZ	TAZ	Units	Comment
Nursing Homes	207	5	20	in Cave Creek
	316	1172	400	in Chandler: at NW corner of Arizona/Chandler
	n/a	235,265	n/a	in Fountain Hills
	n/a	n/a	n/a	in Gilbert
	234	178	100	in Surprise: at 12215 W. Bell Rd.
School Dormitories	n/a	n/a	n/a	in Avondale for Estrella Mountain Community College: on Dysart Rd. just south of boundary with Litchfield Park
	n/a	n/a	n/a	Williams AFB reuse possible
	n/a	662	n/a	Rio Salado Community College at Bush/McKellips
	n/a	421	n/a	in Paradise Valley for Judson School: at Indian Bend Rd/ Mockingbird Lane
Jails	239	306	160	in Peoria: NW Maricopa County Complex at 8401 W. Monroe

Table 6-12
Allocation of Dormitory Populations
(from ESG Final Report)

Year	Total	ASU	%	Grand Canyon	%	Williams ReUse	%
1993	4,242	3,661	0.8630	581	0.1370	0	0.0000
1995	4,868	4,170	0.8566	594	0.1220	104	0.0214
2000	5,487	4,545	0.8283	628	0.1145	314	0.0572
2015	8,509	7,246	0.8516	856	0.1006	407	0.0478
2040	12,702	10,817	0.8516	1,278	0.1006	607	0.0478

6.2.4 Land Absorption

ESG also prepared estimates of land absorption rates for each of the land uses which are consumed by *special population groups*. These were as follows:

- *Motels/hotels/resorts:* Since 1984 aggregate absorption rates for motels, hotels, and resorts has run at 43.97 rooms per acre, with occupancy decreasing from 1.39 persons per room in 1990 to 1.28 persons per room in the year 2040.

- *Mobile Home Parks:* Since 1984 absorption rates for mobile home parks has averaged 6.63 units per acre.

- *RV Parks:* Since 1984 absorption rates for RV Parks has averaged 12.84 spaces per acre.

- *Jails:* Outlying prison facilities average 15.56 persons per acre. Downtown Phoenix prison facilities run 442.80 persons per acre.

- *Dormitories:* Dormitories at ASU average 44 rooms per acre. With average occupancy of 1.5 persons per room, this is equivalent to 66 students per acre.

- *Nursing Homes:* Since 1984 nursing homes have averaged 27.6 beds per acre.

6.3 The SubArea Allocation Model

In this section the requirements of the Subarea Allocation Model (SAM) is described. The objective of the SAM is to disaggregate forecast growth in land uses associated with the special population groups into small grid cells by location; that is to generate forecasts of military housing, dormitories, jails, nursing homes, mobile home parks, RV parks, retirement communities, and hotels/motels/resorts at a microscopic-scale geography. While MAG's existing methodology for generating countywide forecasts of special population groups will be used, the SAM will be responsible for disaggregating these countywide control totals to subareas.

Once done, these coverages can be used to generate estimates of special population groups on any geography through simple aggregation.

A full discussion of the SAM will be given in Chapter 8. The salient elements applicable to the analysis here are as follows:

- *Existing land use covers* must be generated for a base year which depict locations and attributes associated with special population land uses.

- Data which will enable us to *predict* the location and extent of land use expansion associated with *growth* in special populations needs to be developed.

- Methods for *converting* between special population groups (seasonal, transient, and group quarters populations) and their land use counterparts (motels, mobile home parks, nursing homes, etc.) must be developed.

6.4 Seasonal Populations

Conversion

As discussed in Section 6.2, overall seasonal population forecasts are based on three independent forecasts of seasonal populations in (1) mobile homes, (2) retirement communities, and (3) RV parks. It will be necessary to *preserve* each component in order to translate seasonal population estimates into the respective number of housing units associated with each. (That is to say, a single overall control total for seasonal population does not give us enough information to estimate the proportion in mobile homes, the proportion in retirement communities, and the proportion in RV parks. These breakdowns were created from the original forecasting process).

Table 6-13 gives factors to enable us to do this. These were merely derived from data already presented earlier from the ESG report. Note that these data already take into consideration occupancy rates, expansion not related to seasonal populations, etc.

Table 6-13
Conversion Factors for Seasonal Populations

Retirement Communities	1.286 persons/dwelling unit	From Table 6-3
Mobile Homes	0.783 persons/dwelling unit	From Table 6-2
RV Parks	1.476 persons/RV space	From Table 6-4

These factors will enable us to estimate growth in mobile home spaces (for example) throughout Maricopa County from growth in seasonal population. The SAM will have to identify locations where this growth could be expected to occur. Similarly, these factors will enable us to estimate seasonal populations for any zone system (say TAZs) from a land use cover that contains information about the location of mobile home parks (both now and in the future).

Database Structure

ESG created an inventory of mobile home and RV parks in 1994, including geocoding addresses associated with these locations. A *point coverage* of these locations is built to use as a basis for the SAM (which will be rasterized to a grid cell structure). Existing parks grided from the point coverage will give us an ability to generate estimates of seasonal populations for any zone system (e.g., TAZs).

The point coverage database structure is shown in Table 6-14. Only the pertinent information needed is extracted by the SAM from the original ESG inventory. Note that ESG provided estimates of overall parcel size in their inventory. The coverage, to describe existing 1993 conditions, has been installed at MAG at ***/magtpo/sam/base/mobileh.***

Table 6-14
Mobile Home/RV Park Coverage

Field	Datatype	Description
mobileh #	binary integer	Internal ARC/INFO identifier
mobileh-id	binary integer	Internal ARC/INFO identifier
mhpname	character	Name of establishment
totspaces	integer	Total MH and RV spaces on site
rvspaces	integer	Number of RV spaces
mhspaces	integer	Number of MH spaces
landsize	decimal	Parcel size in acres TAZ = integer RAZ = integer
type	character	MH = mobile homes only RV = rv spaces only B = both

A similar coverage to describe retirement community locations is needed, such as those listed earlier in Table 6-6. This coverage does not presently exist, but should be created. Nothing more than polygons is envisioned which outline parts of Maricopa County that are eligible for development to support that component of seasonal population

Projection Methodology

For future projections, the ability to allocate growth in seasonal populations is needed (by land use type) to locations throughout Maricopa County. In ARC/INFO grid, this will mean that individual candidate grid cells need to be "selected" for development. In order to do this, (1) land absorption rates associated with each of these types of land use and (2) typical project sizes.

The idea behind the second point, typical project sizes, is that mobile home park expansion does not occur a single unit at a time. Instead, individual mobile home park sites are built with some typical range of capacity. Growth in the mobile home park component of seasonal population should reflect this. Figure 6-1 illustrates the distribution of mobile home (and RV) park sizes currently reflected in the ESG inventory. Table 6-15 shows average park sizes and land absorption rates associated with them, both overall as well as for just those parks developed in the last 10 years (since 1984). (Note that contemporary parks are considerably larger than those built more than 10 years ago).

Figure 6-1

Distribution of Mobile Home Park Sizes

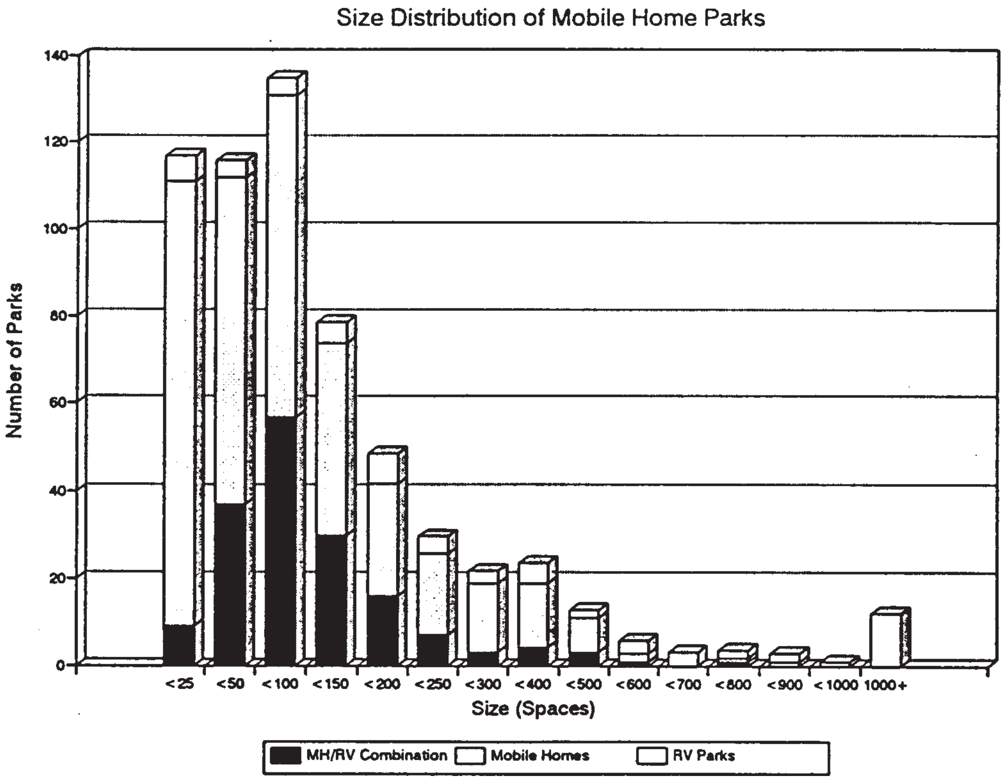


Table 6-15
Average Sizes Associated with Mobile Home and RV Parks

		Number Sites	Total Spaces	Total Acres *(millions)*	Avg Size *(spaces)*	Spaces/ Acre
All	MH Parks	383	39,257	5,072.5	102	7.7
	RV Parks	64	31,583	2,653.1	493	11.9
	Combination	169	19,646	1,687.0	116	11.6
	Overall	**616**	**90,486**	**9,412.6**	**147**	**9.6**
Post 1984	MH Parks	20	5,012	1,215.3	251	4.1
	RV Parks	10	9,770	1,175.9	977	8.3
	Combination	2	909	75.0	455	12.1
	Overall	**32**	**15,691**	**2,466.1**	**490**	**6.4**

Table 6-15 implies that, on the average, mobile home parks are built with a capacity of 251 units, consuming 61 acres of land (based on post 1984 developments). *(Note that these land absorption rates are close to what ESG reported in their final report.)*

Locational Preferences

Figure 6-2 illustrates the locations of mobile home/RV parks throughout Maricopa County established by the inventory created by ESG. Note, however, that this coverage is not a good reflection of where growth in this land use sector is likely to occur. A better image of growth potential is shown in Figure 6-3, which illustrates the locations of those parks which have been built since 1984. The locational characteristics evident are:

- Locational preferences of mobile home parks and RV parks appear to be generally the same.
- Growth is occuring in distant outlying areas (such as Apache Junction). This probably reflects underlying zoning requirements and the level of acceptance of these types of developments in various communities
- Growth is oriented primarily to general freeway corridor locations

These locational preferences can be implemented in the SAM by building a polygon coverage that establishes zones where mobile homes will be located.

In addition, it is necessary to use underlying general plan information to help identify appropriate locations for MH/RV parks. Inspection of the general plan coverage reveals, however, that there is only one general plan code (37) for RV parks — but no instances of its appearance in the general plan cover. There are no codes specifically associated with MH parks.

Figure 6-2
Mobile Home/RV Park Locations in Maricopa County

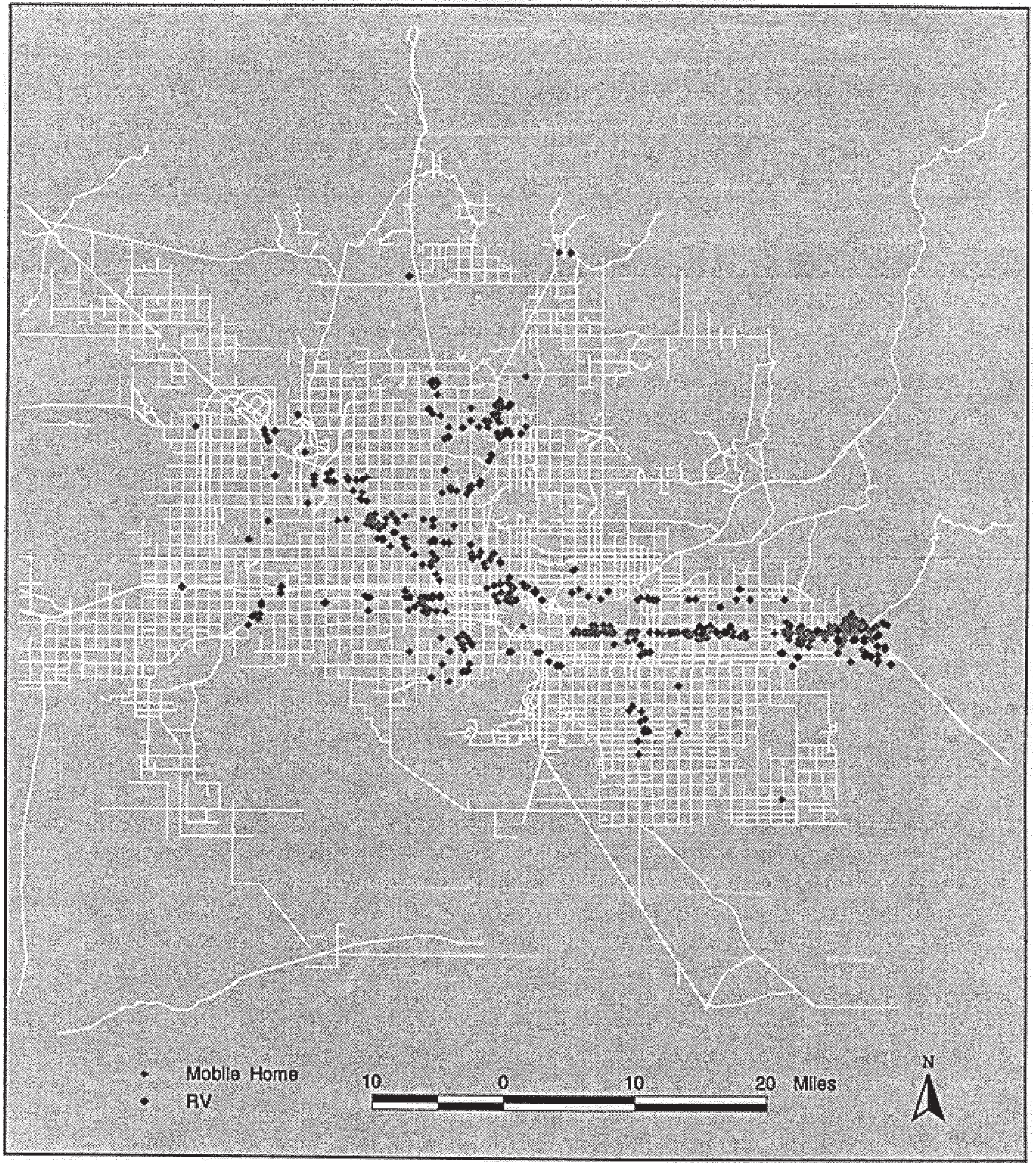
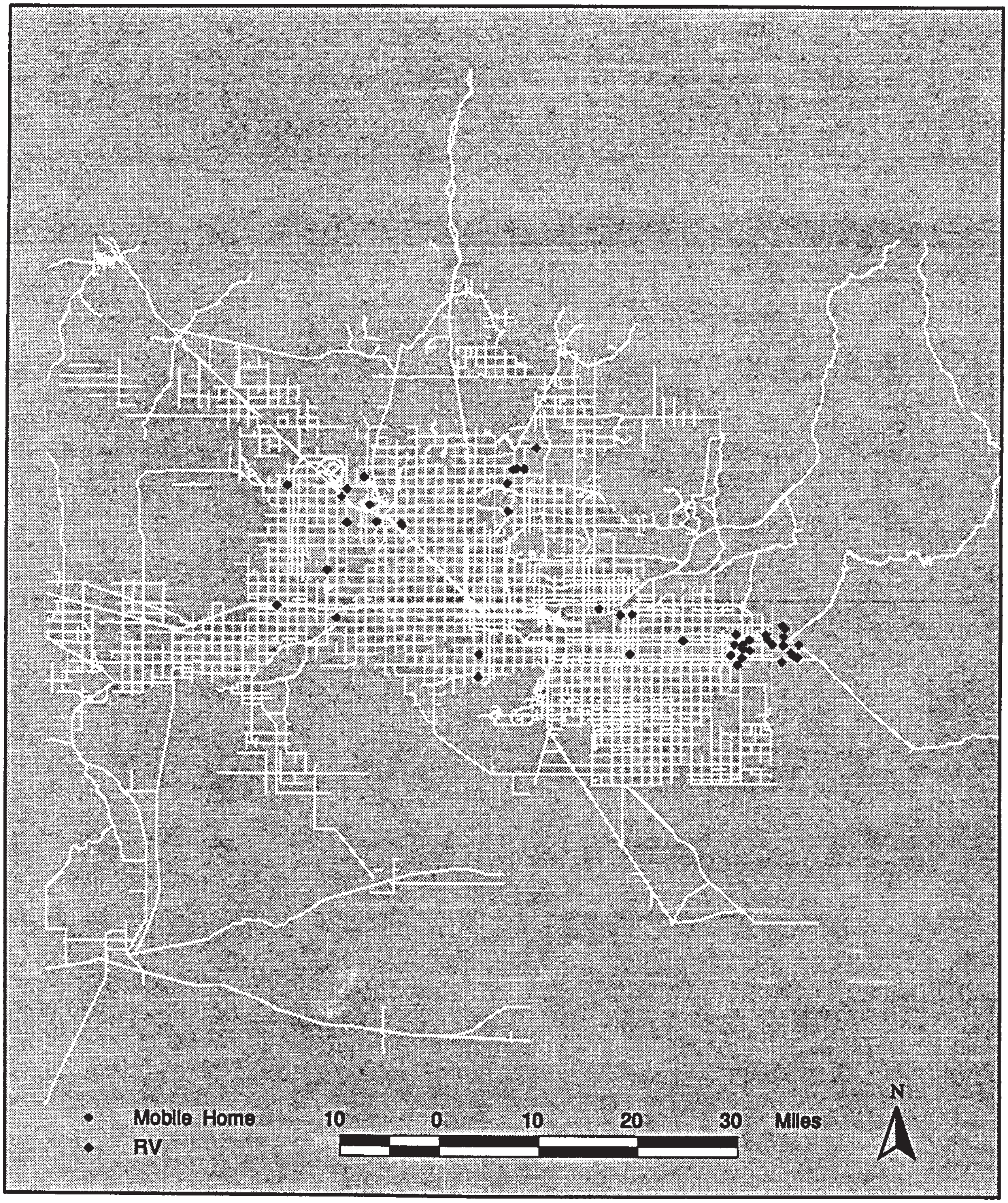


Figure 6-3

Mobile Home/RV Park Locations in Maricopa County Built Since 1984



Consequently, the existing land use cover in order to get some idea about typical land uses associated with MH/RV locations is also examined. This subject is covered next.

Consistency with the Existing Land Use Cover

A point-in-polygon analysis to compare the MH/RV park coverage with the existing land use cover is performed. As can be seen in Table 6-16, the existing land use cover only partially reflects the existence of MH/RV parks around the region with 37% of the locations falling into a land use polygon that properly reflects mobile homes.

Table 6-16
MH/RV Park Conformance with the Existing Land Use Cover

Existing LU Code	Description	Number of Parks	%
0	Unknown	43	7.0
1-3,5	Residential	125	20.3
4	Mobile Home Parks	229	37.2
6-7	Commercial	95	15.4
8	Hotel	4	0.7
11-12	Office	2	0.3
13-14	Industrial	23	3.7
21	Public	3	0.5
22	churches	4	0.7
24-26	Rights-of-Way	8	1.3
27	Parks	3	0.5
28	Golf Courses	1	0.2
29-30,36	Waterways	1	0.2
31	Vacant	64	10.4
33	Agriculture	9	1.5
40	NonDevelopable	1	0.2
	Total	616	100.0%

The fact that the existing land use cover does not seem to reflect the existence of MH/RV parks in many instances is somewhat disturbing — it should be improved in the future. Still, it will not affect the performance of the SAM. The ESG inventory as accurate and will override any conflicting information carried in the existing land use cover is merely accepted.

What the conflict seems to suggest, however, is that residential (presumably single family lower density) and commercial land uses are most often associated with MH/RV parks. These will be the general plan land use codes that will be considered to be eligible for mobile home/RV park expansion (in addition to areas of the county for which there is no general plan designation).

6.5 Transient Populations: Hotels, Motels, and Resorts

Conversion

As discussed in Section 6.2, three land use components together comprise overall transient population: (1) motels, (2) hotels, and (3) resorts. Forecasts are predicated on an assumption that occupancy rates in Maricopa County will gradually increase through the year 2040, therefore the factor needed to convert between M/H/R rooms and transient population changes by year. Table 6-17 reports these factors, which were merely derived from data already presented earlier from the ESG report.

Table 6-17
Conversion Factors for Transient Populations

	Transient Population per Room		
Year	Motels	Hotels	Resorts
1990	0.825	0.825	0.825
1995	0.832	0.832	0.832
2000	0.834	0.834	0.834
2015	0.859	0.859	0.859
2040	0.898	0.898	0.898

Visitors per room rates are the same for motels, hotels, and resorts
Data was derived from Table 6-8 (column 5) / Table 6-9 (column 7)

These factors will enable us to estimate growth in hotel space throughout Maricopa County from growth in transient population. The SAM will have to identify locations where this growth could be expected to occur. Once M/H/R land use (current or forecast) is represented in disaggregate form, these same factors can be used to generate estimates of the transient population associated with it for any zone system (say TAZs).

Database Structure

ESG created an inventory of hotel, motel, and resort space in 1994, including geocoding addresses associated with their locations. A *point coverage* of these locations to use as a basis for the SAM is built (which will be rasterized to a grid cell structure). Existing hotels gridded from the point coverage will give us an ability to generate estimates of transient populations for any zone system (e.g., TAZs).

The point coverage database structure is shown in Table 6-18. Only the pertinent information needed by the SAM from the original ESG inventory is extracted. However, note that ESG *did not* provide estimates of overall parcel size in their inventory.

Table 6-18

Motel, Hotel, Resort Coverage

Field	Datatype	Description
hotel #	binary integer	Internal ARC/INFO identifier
hotel-id	binary integer	Internal ARC/INFO identifier
rooms	integer	Number of rooms
type	character	M=motel, as coded originally by ESG H=hotel R=resort
bldg name	character	Name of establishment
taz	integer	TAZ number
raz	integer	RAZ number

The coverage, describe existing 1995 conditions, has been installed at MAG at ***/magtpo/sam/base/hotels.***

Land Absorption

As was mentioned earlier, the ESG inventory of M/H/R space did not include estimates of parcel size which is needed in order to estimate land absorption associated with M/H/R facilities of various sizes (measured by the number of rooms). Two different types of analyses were attempted to obtain information on parcel sizes:

- Hotel, motel, and resort locations were overlaid on the existing land use coverage in hopes to derive land absorption rate from the areas of land use polygons contained in that cover. Unfortunately, it was found that the existing land use cover usually does not reflect hotel and motel locations (in over 90% of the cases). Actual land uses recorded in the existing land use cover are shown in Table 6-19.

  Because it was clear that polygons in the existing land use cover included many uses besides just hotels, it could not be used as a basis to determine land absorption rates.

- The M/H/R dataset also contained the parcel number (the APN) for the property, and therefore parcel sizes from the parcel inventory could be extracted from the parcel coverage. The results of this analysis are shown in Table 6-20. This operation appeared to work effectively — about 75% of the parcels matched were designated as hotel sites in the parcel file (the other 25% were designated multiple family).

The results of this analysis show that motels, hotels, and resorts have run at a gross density level of 29.6 rooms per acre, since 1984. In Section 6.2 it is noted that ESG reported an overall density level of 43.97 rooms per acre which is considerably higher; however it is not

Table 6-19

Hotel Conformance with the Existing Land Use Cover

Existing LU Code	Description	Number of Hotels	%
0	Unknown	12	n/a
1-4	Residential	47	13.7
6-7	Commercial	153	44.6
8	**Hotel**	**32**	**9.3**
9	Mall	5	1.5
11-12	Office	10	2.9
13-14	Industrial	27	7.9
21	Public	1	0.3
24-26	Rights-of-Way	8	2.4
27	Parks	3	0.9
28	Golf Courses	4	1.2
29-30,36	Waterways	3	0.9
31	Vacant	47	13.7
33	Agriculture	3	0.9
	Total	**355**	**100.0%**

Table 6-20

Land Absorption and Size Characteristics for Hotels

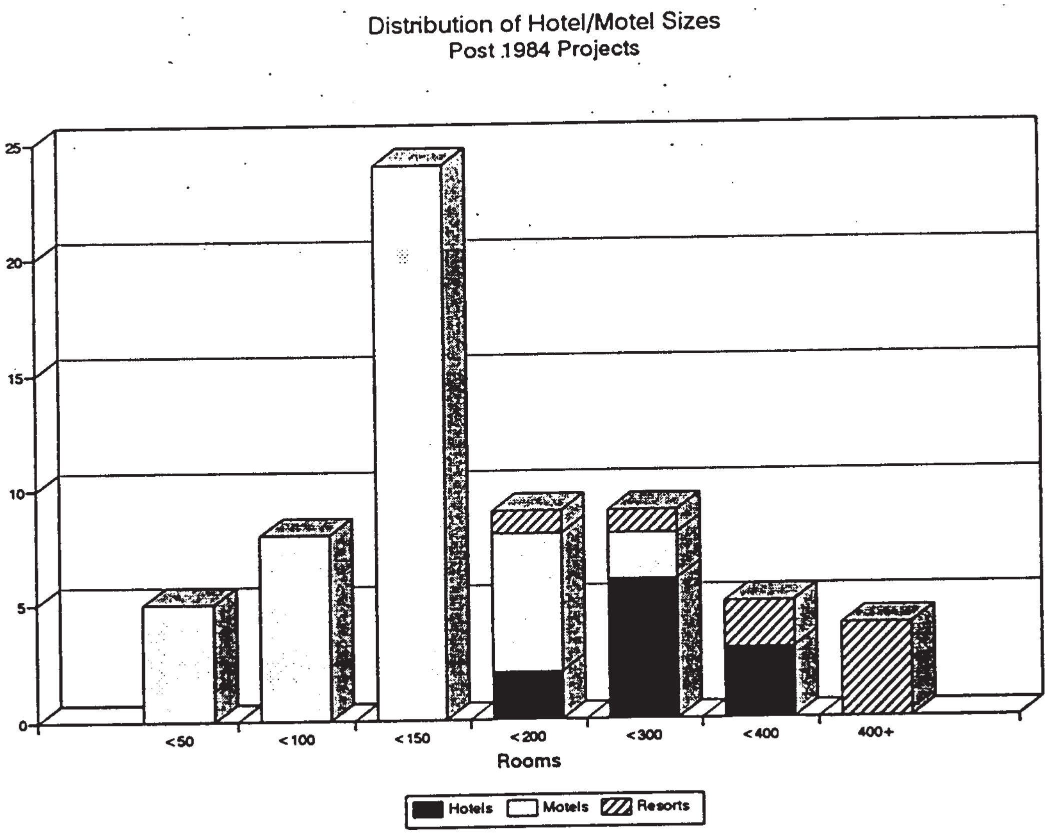
		Number Sites	Number Rooms	Parcel Acres	Average Size	Rooms/Acre
Overall	Hotels	15	3,698	97.5	247	37.9
	Motels	214	18,191	497.5	85	36.6
	Resorts	22	7,311	842.7	332	8.7
	Overall	**251**	**29,200**	**1,437.7**	**116**	**20.3**
Post 1984	Hotels	11	2,768	68.6	252	40.4
	Motels	37	4,645	101.0	126	46.4
	Resorts	7	2,773	321.3	396	8.6
	Overall	**55**	**10,186**	**490.9**	**185**	**20.8**

Note: Data reported only on those sites which successfully matched the parcel file so average parcel size could be computed.

determine the source of ESG's estimate (since parcel size data does not appear in their hotel coverage). As for mobile home parks, information is also needed about typical motel, hotel, and resort sizes. The distribution of hotel sizes is shown in Figure 6-4.

As noted in Table 6-20, motels, hotel, and resort facilities seem to be growing larger in recent years; since 1984 the average M/H/R facility contains 185 rooms and therefore would absorb 8.9 acres. It is interesting to note that densities for motels and hotels are about the

Figure 6-4
Distribution of Motel, Hotel, and Resort Sizes



same, implying that the distinction between the two in the ESG database is somewhat arbitrary, at least with respect to this characteristic. A more important characteristic with respect to density is location: Motels and hotels located in the CBD together average 473 rooms per acre; elsewhere throughout the region they average 49.8 rooms per acre. (These results are similar to those reported by ESG). This is shown in Table 6-21.

Table 6-21
Average Density Levels for Motels and Hotels Based on Location

Subarea	Total Rooms	Total Acres	Rooms/Acre
Downtown	4,475	9.46	473.0
Airport	4,297	78.27	54.9
Elsewhere	20,881	507.26	41.2
Overall	**29,653**	**595.0**	**49.8**

Note: Resorts were not included in this analysis.

As would be expected, one of the distinguishing characteristics between motels, hotels, and resorts concerns project size: resorts typically are larger facilities than the other two components serving transient populations.

Locational Considerations

Figure 6-5 illustrates the locations of motels, hotels, and resorts throughout Maricopa County represented in the inventory created by ESG. The pattern of growth demonstrated since 1984 shown in Figure 6-6 generally echoes the overall pattern which has been established historically in the Phoenix metropolitan area, although there are some differences. For example, newer facilities tend to be located at more outward locations. Also, the Grand Avenue corridor which figured strongly in early Phoenix hotel development no longer supports this type of development activity at all.

The observations about general locational aspects of hotels, motels, and resorts, are as follows:

- Location preferences applicable to resorts are different than for hotels and motels. They do not appear to be dependent on drive-by traffic or local concentrations of activity. Since they require large tracts of land, they are typically more remotely located. In Maricopa County, they are heavily concentrated in the northeast quadrant of the metropolitan area.

- There seem to be three principal markets for motels/hotels: (1) downtown Phoenix, (2) Phoenix airport, and (3) along freeway corridors.

Figure 6-5
Motel, Hotel, and Resort Locations in Maricopa County

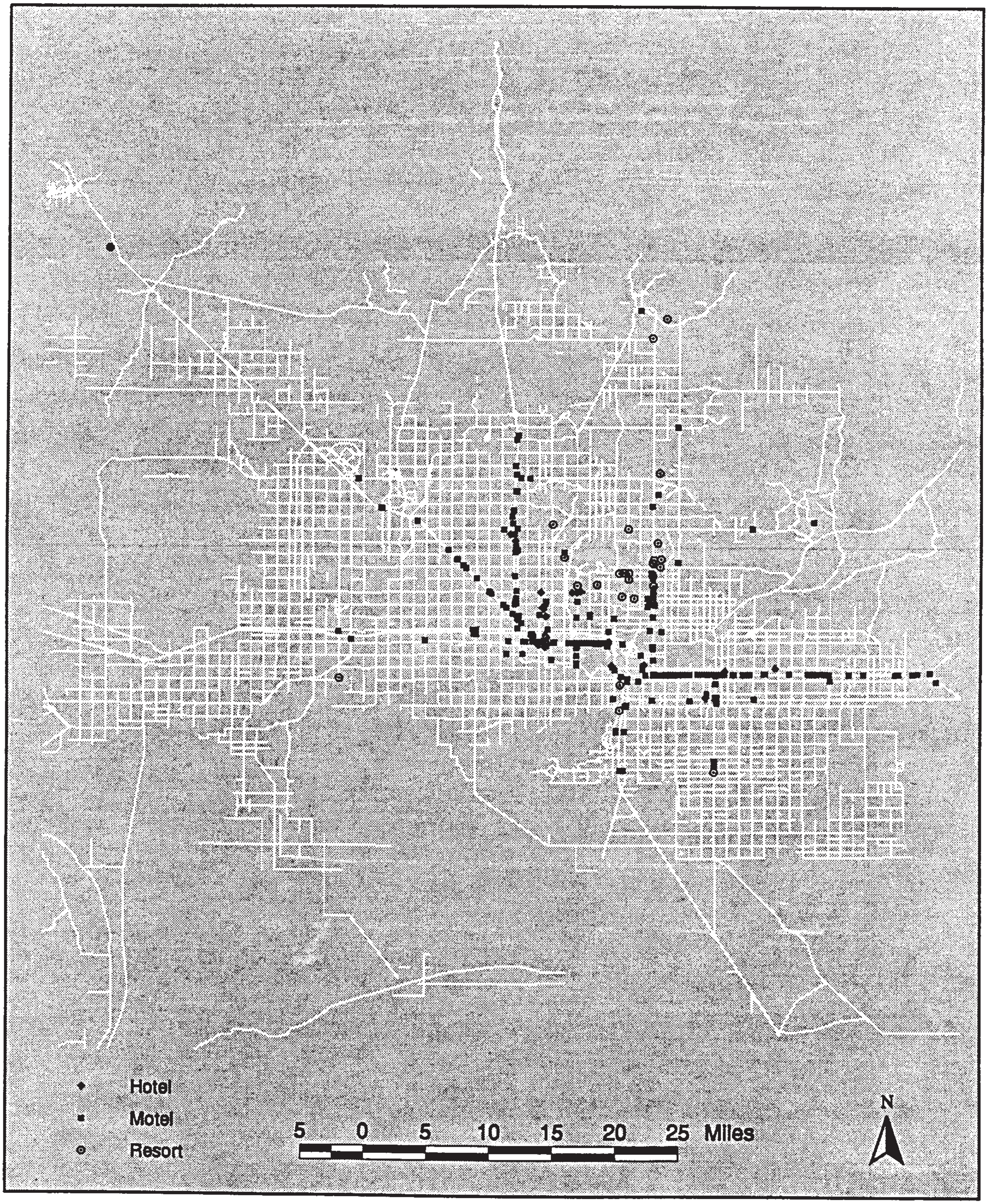
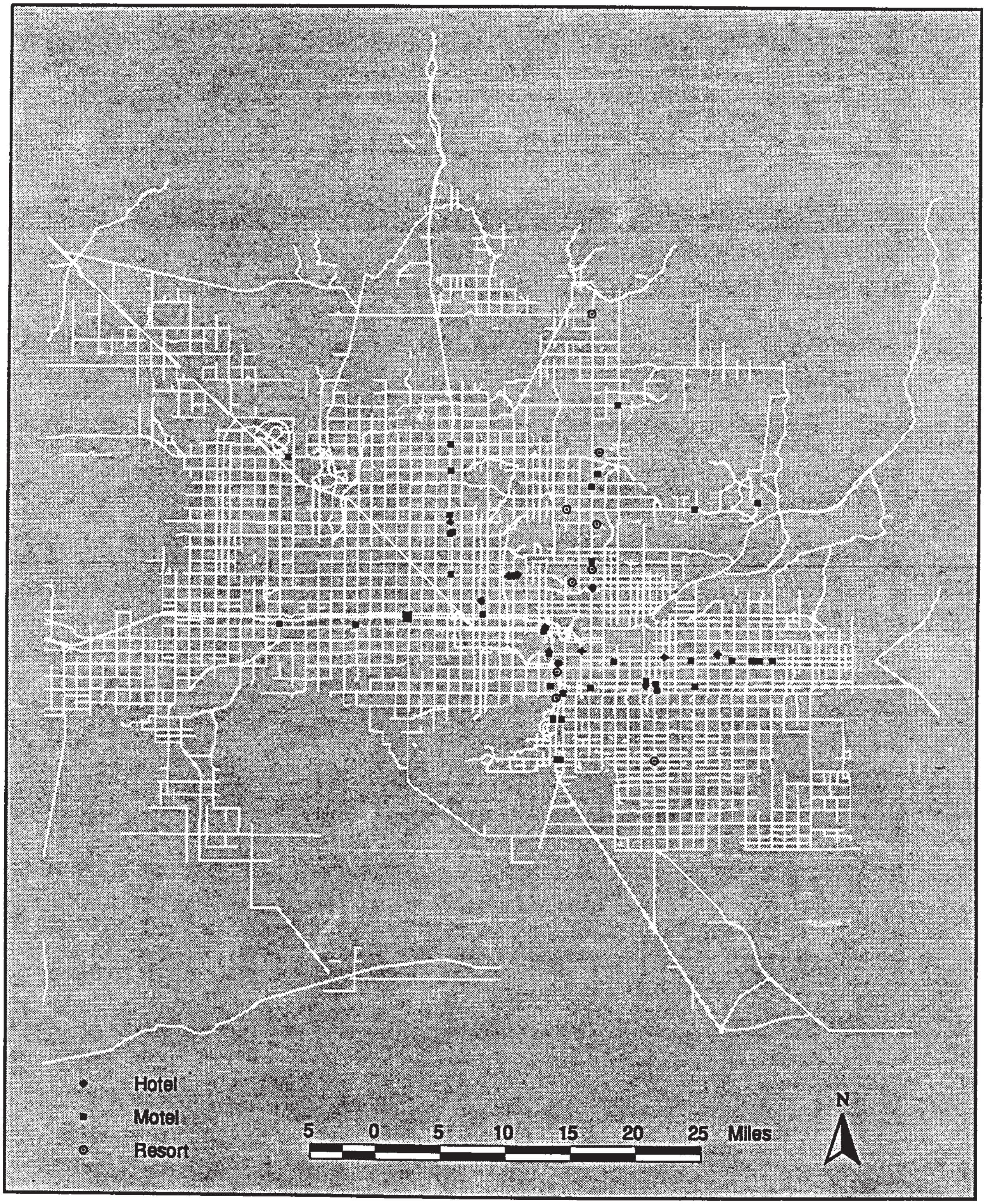


Figure 6-6
Motel, Hotel, and Resort Locations in Maricopa County Built Since 1984



It would seem then that motel and hotel growth should be allocated separately from resorts. Table 6-22 gives a breakdown of these markets, and shows that each market seems to have remained stable over time with the exception of downtown, which has seen little development in the last 10 years.

Table 6-22
Breakdown of Motel, Hotel, and Resort Development Trends

	Subarea	Total Rooms	%
Overall	Airport	4,297	11.1
	CBD	4,475	11.6
	Elsewhere	8,406	55.4
	Resorts	8,406	21.8
	Total	**38,555**	**100.0**
Since 1984	Airport	1,160	9.0
	CBD	273	2.1
	Elsewhere	7,714	60.2
	Resorts	3,674	28.7
	Total	**12,821**	**100.0**

In addition, underlying general plan information is wanted to help identify appropriate locations for M/H/R development. Inspection of the general plan coverage is revealing:

- There are three codes (45, 153, and 206) which identify hotel/motel uses, however there are no instances in the general plan cover where they are used.
- There are seven codes (55-58, 201, and 207) which identify resort uses, and there are 29 areas in the general plan identified for resorts (possibly this reflects existing uses also).

As noted earlier, motel/hotel development is most frequently associated with general commercial land uses (in the existing land use cover); it is reasonable to consider retail and general commercial uses in the general plan cover to also be considered eligible for motel and hotel development.

6.6 Group Quarters

Group quarters populations include residents of (1) nursing homes, (2) jails and prisons, (3) military housing, and (4) dormitories. With the exception of nursing homes, there are very few instances of these land uses:

- No expansion in military housing is reported in the ESG projections.
- Minor expansion in dormitory housing is anticipated, and then it will naturally fall immediately on campuses for the colleges/universities in question.
- Expansion in prisons and jails necessary to meet future requirements is best handled externally from the SAM — this is a public policy decision and there is no market driven forces that influences the locations of prisons and jails.

All of these types of institutions are in the coverage ***/magtpo/sam/base/inst***. A description of the database is in Table 6-23.

Table 6-23
Other Group Quarter Coverage

Field	Datatype	Description
inst #	binary integer	Internal ARC/INFO identifier
inst-id	binary integer	Internal ARC/INFO identifier
propname	character	Number of institution
beds	integer	Number of beds
pop	integer	Institutional population
taz	integer	TAZ number
raz	integer	RAZ number
type	character	type of institution p = prism j = Judson school dormitory o = other school dormitory a = ASU dormitory g = Grand Canyon college dormitory

Nursing homes and other similar congregate care facilities, however, are different. These will be discussed next.

Conversion

ESG developed a database of nursing home facilities during the 1994 study, along with geocoded coordinates for their locations. This too was built into an ARC/INFO point coverage and has been installed as ***/magtpo/sam/base/nursing***. The total capacity of nursing home facilities contained in this database is 23,897 beds, so an appropriate conversion factor between nursing home capacity and the associated group quarter population (from Table 6-10) would be 0.397 persons per bed which itself reflects the occupancy rate of nursing home facilities for permanent residents.

Database Structure

The database structure for nursing homes is shown in Table 6-24.

Table 6-24
Nursing Home Coverage

Field	Datatype	Description
nurse #	binary integer	Point cover internal identifier
nurse-id	binary integer	Point cover internal identifier
company	character	Name of establishment
beds	integer	Number of beds
taz	integer	TAZ number
raz	integer	RAZ number

Land Absorption

ESG reported gross density levels of 27.6 beds per acre for nursing homes. This data was attempted to be verified; however there is no information in the nursing home database about parcel size nor was there any information on which to base an analysis (unlike hotels, there was no parcel number either). Consequently, the ESG analysis was accepted.[2]

Note that the areas associated with sites appearing in the database will have to be estimated based on this density level.

The distribution of nursing home sizes reflected in the database is shown in Table 6-8. The average of a nursing home facility is 61 beds, however the distribution shown in Figure 6-7 clearly indicates nursing homes are typically very small — 54% consist of 20 beds or fewer.

Locational Considerations

Figure 6-8 illustrates the locations of nursing homes in Maricopa County. As is evident in the graphic, nursing homes are generally uniformally distributed geographically, but only in the northeastern parts of the region. The southwest is remarkably void of any instances of nursing home facilities. ESG also noted in their report that nursing homes favored hospital locations. This in any type of location model would be avoided, however, because it would only introduce the need to forecast hospital locations in the future.

[2] It appears reasonable though — this density level is about the same as that determined for motels assuming that motels typically offer two beds per room.

Figure 6-7
Distribution of Nursing Home Sizes

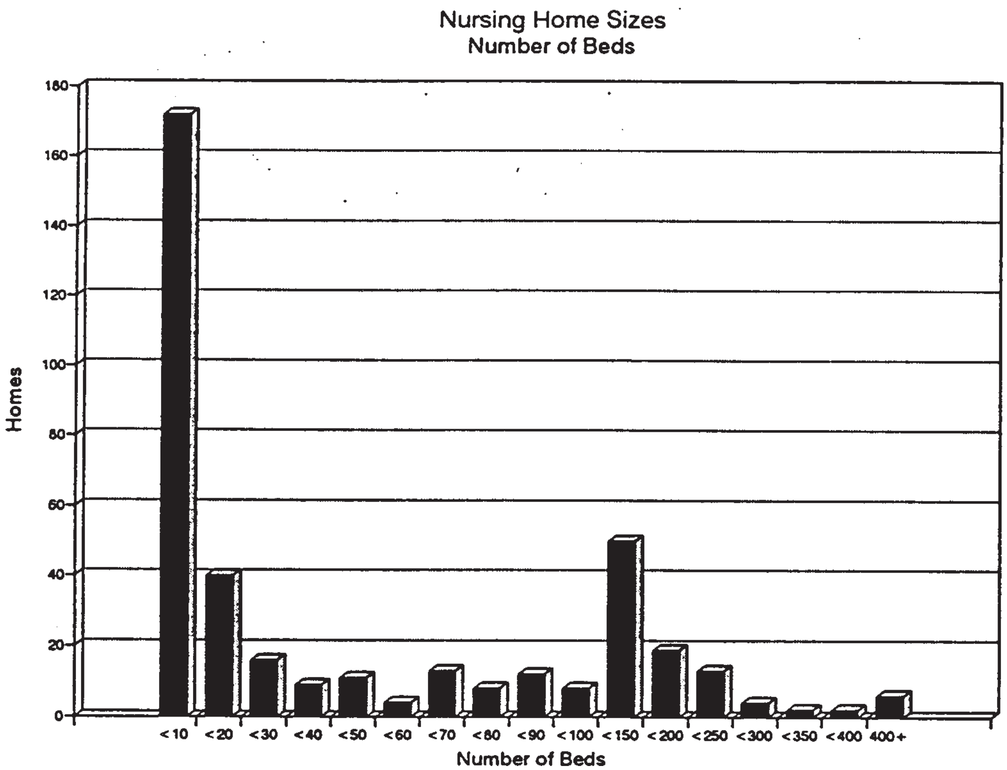
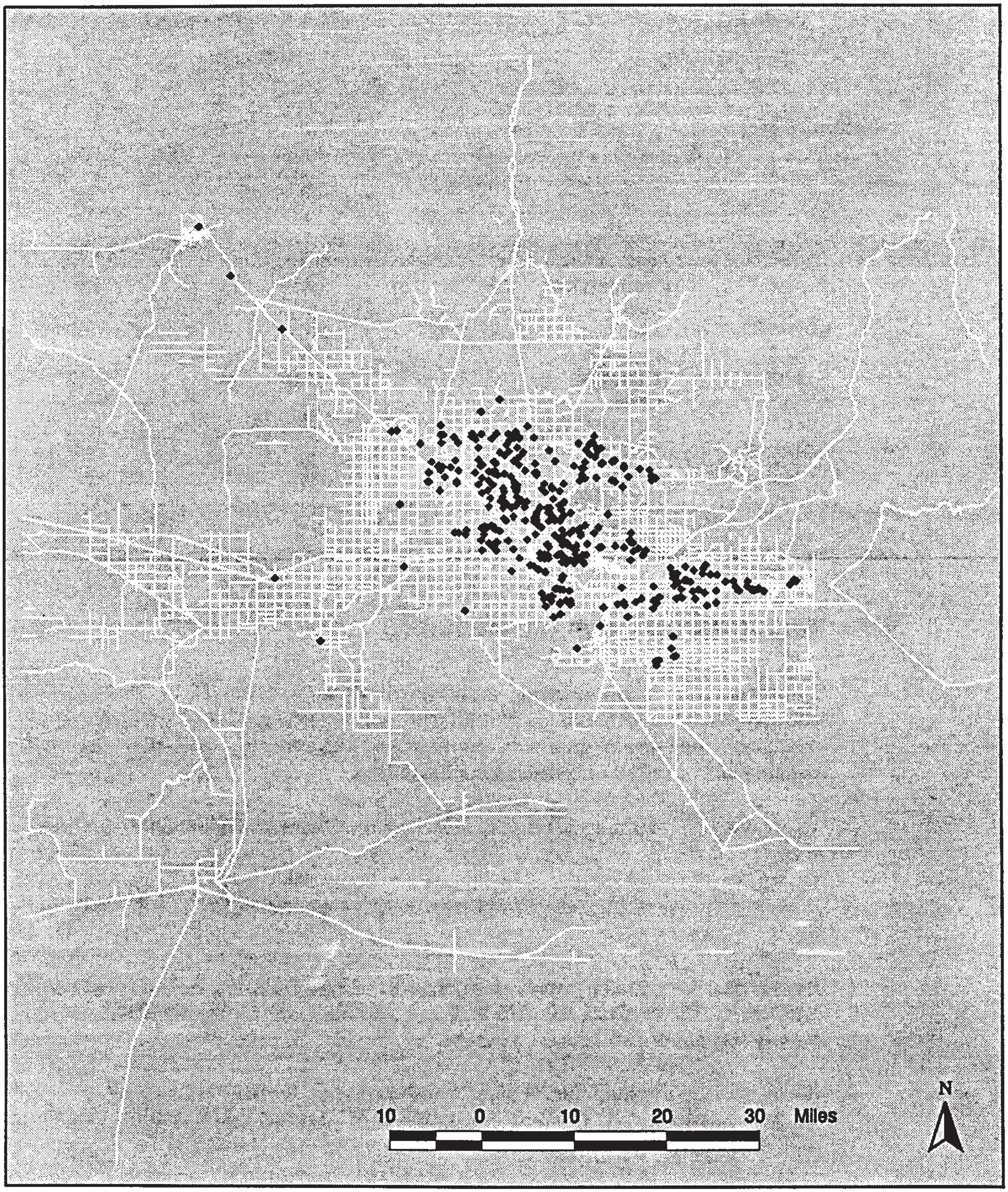


Figure 6-8
Nursing Home Locations in Maricopa County



There are no general plan codes which designate lands for nursing homes. An overlay of nursing home locations with the existing land use cover (see Table 6-25), however, shows that residential, hospital, and commercial lands are those where nursing homes are typically found. Therefore, these types of general plan codes will be used as candidate locations in the SAM model.

Table 6-25

Nursing Home Conformance with the Existing Land Use Cover

Existing LU Code	Description	Number of Hotels	%
0	Unknown	11	2.7
1-4	**Residential**	**266**	**64.4**
6-7	**Commercial**	**53**	**12.8**
8	Hotel	1	0.2
11-12	Office	8	1.9
13-14	Industrial	6	1.4
19,20	**Hospitals**	**28**	**5.8**
21	Public	1	0.2
24-26	Rights-of-Way	2	0.5
29-30,36	Waterways	2	0.5
31	Vacant	22	5.3
33	Agriculture	9	2.2
	Total	**413**	**100.0**

6.7 Recommendations for the Future

1. Improvements to the existing land use cover could be made in the future to assure consistency between it and the point coverages describing seasonal, transient, and group quarters populations. Currently, there appears to be considerable conflict between these databases.

2. Coverages of known projects listed in various tables in this section need to be made.

3. A coverage describing retirement communities shown in Table 6-6 needs to be made.

4. Research into average parcel sizes connected with nursing homes needs to be conducted. The analysis failed to find any reliable existing source of information among the MAG datasets.

5. Research into average parcel sizes connected with institutional group quarters needs to be conducted. Currently, existing MAG datasets do not allow this to be analyzed either.

7
DRAM/EMPAL Feedback Mechanism

Along with a number of other metropolitan planning organizations around the country, MAG relies on DRAM/EMPAL to generate forecasts of land use throughout Maricopa County. These land use forecasts are used to create trip generation datasets which are used by the travel demand forecasting process, which in turn are used to generate estimates of vehicle emissions and air quality. Therefore, the DRAM/EMPAL land use forecasting model plays an important role in the fulfillment of MAG's mission.

The DRAM/EMPAL software is used at MAG to generate forecasts for RAZs (which are aggregations of TAZs). DRAM generates demographic forecasts for five household classes (quintiles) of income. EMPAL generates forecasts for five categories of employment (office, retail, industrial, government, and other). DRAM/EMPAL historically has been operated on personal computers. In addition to the model itself, the MAG process has involved a complicated program of DOS and Paradox scripts to manage the various databases which feed data to DRAM/EMPAL and to manage its results.

One of the concepts integral to the model's operation is the notion that growth patterns in regions are a function of accessibility. Accessibility measures are one of the key inputs to the modeling process and are expressed in terms of travel times between RAZs. Travel times between representative TAZs, one for each RAZ, are typically pulled from the travel model output from a previous forecast year, and used as input for the new forecast year.

One of the issues which has arisen at MAG (and with DRAM/EMPAL users in other metropolitan areas, as well) concerns whether the resulting land use forecasts are consistent with the accessibility assumptions on which they are based. This raises the question as to whether travel times between RAZs resulting from a land use forecast would be similar to the travel times which were assumed.

The solution to this question has led to the concept of a *feedback loop* — this is, intermediate land use forecasts are fed into the travel demand modeling process to generate new estimates of travel times which are used by DRAM/EMPAL for a subsequent iteration. This is a cyclical process which, if it closes on a unique solution, would tend to attenuate travel times so that DRAM/EMPAL arrives at a distribution consistent with the travel times on which it is based.

7.1 Objectives and Background

The objective of this task was to implement a feedback loop mechanism for DRAM/EMPAL.

MAG recently completed the conversion of its travel demand forecasting process to the EMME/2 software package operating on a unix-workstation. In addition, MAG has upgraded its DRAM/EMPAL license to operate on a unix-workstation. Therefore, the approach involved a *re-implementation* of the land use forecasting process on unix-work stations, with a feedback loop mechanism. MAG's GIS ARC/INFO was used as the front-end of the process, replacing the DOS and Paradox scripts that MAG formerly used in a personal computer environment. Therefore, the work involved much more than simply *porting* programs from the DOS environment to the unix environment — the code was entirely rewritten.

Inner and Outer Loops

There is extensive documentation available on the theory and operation of DRAM/EMPAL and will not be repeated here. A brief overview of the processing steps, however, would be appropriate:

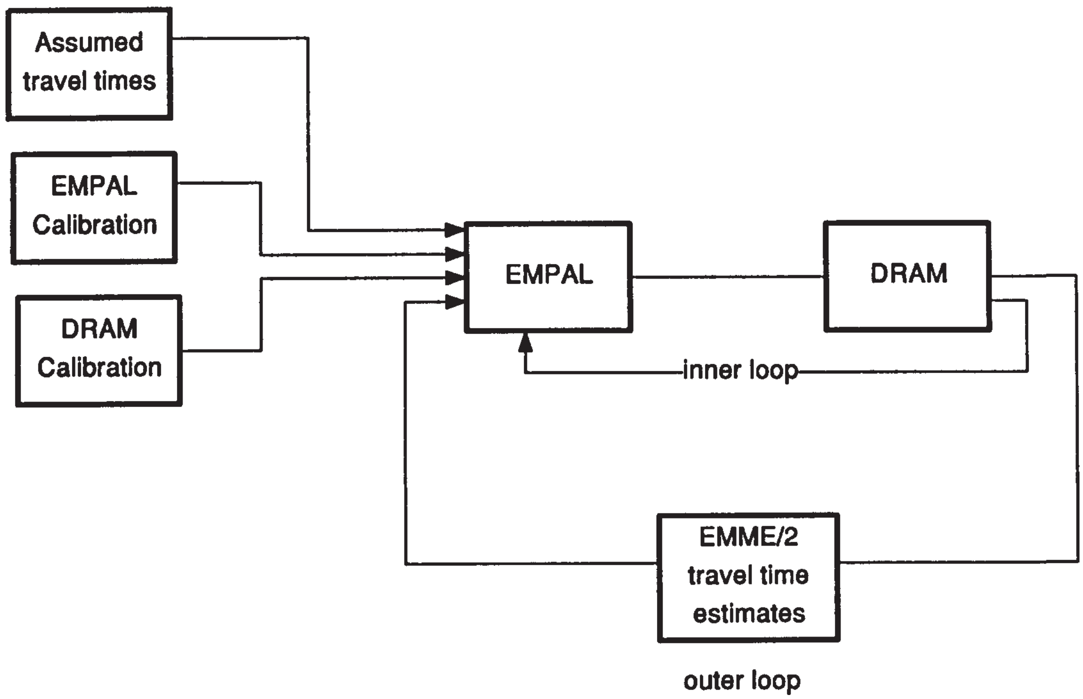
- DRAM/EMPAL is first calibrated for a base year. This calibration is performed for the purposes of establishing so-called k-factors which are used for future forecasts.

- For any given forecast year, EMPAL is run first to provide projections of employment by five employment classes for RAZs. Among other inputs, EMPAL accepts a RAZ-to-RAZ travel time matrix which provides measures of accessibility. EMPAL also accepts measures of constraints to prevent it from projecting employment levels which are beyond the holding capacity of land available for development within individual RAZs. The EMPAL process iterates cyclicly, adjusting parameters so that land capacity constraints are observed in the final solution.

- DRAM is then run to provide projections of households by income class for RAZs. The concept behind DRAM is similar to EMPAL in that it is operated iteratively, adjusting parameters to observe holding capacity constraints of land available for development within RAZs.

So noted, each DRAM/EMPAL run *itself* is an iterative process — what has become known as the *inner loop*. The objective of this task is to regenerate travel times with MAG's EMME/2 travel model and to repeat the DRAM/EMPAL process with the new set of travel times. This has become known as the *outer loop*. The concept is diagramed in Figure 7-1.

Figure 7-1
The DRAM/EMPAL Feedback Mechanism



Solution Closure

An important issue to address at this point concerns whether the outer loop will reach closure or results will vary wildly from one iteration to another. The latter would result in a forecast solution which would depend almost entirely on the arbitrary decision to halt the process on the *nth* iteration.

Putnam, the architect of DRAM/EMPAL, and Florian, the architect of EMME/2, have researched this question. Putnam argues that the feedback loop mechanism will close on a unique solution if TAZ travel times in *successive* EMME/2 runs are predicated on a weighted average of traffic volumes generated on *previous* EMME/2 runs. The weighted average of forecast volumes should be predicated on a convex function. That is to say, the volume on which travel time computations should be based during iteration n should be:

$$AV_n = \frac{1}{n}V_n + (1 - \frac{1}{n})AV_{n-1}$$

where AV_n is the average volume after iteration n and $_nV$ is the assigned volume from iteration n.

So, subsequent iterations of assignment volumes have decreasing effect. It can be readily shown that this equation reduces to a simple arithmetic average of the results from each previous run:

$$AV_n = \frac{\sum_{i=1..n} V_i}{n}$$

This closure method was utilized in determining the number of iteratives requested for the outer loop.

New Methodology

Generally speaking, the nature of this assignment involved substantial application programming but little new methodological development. Insofar as DRAM/EMPAL is concerned, application programming focused primarily on replacing the implementation of driver and database scripts that formerly worked with DOS and Paradox, but now had to work with unix and INFO. The approach to EMME/2 was similar — in general the attempted to call upon all of the standard EMME/2 macros that are used by MAG's regular travel demand modeling process. Several methodological issues were addressed, however:

- The modeling process does not utilize MAG's new procedures for iterative trip distribution. This new procedure was still under development.

- The modeling process focuses only on P.M. peak period highway travel.[1]

- The mode choice model is not utilized. Since regional mode choice in Maricopa County runs only about 2%, it is assumed that land use decisions respond to highway, rather than transit, accessibility.

- The travel forecasting process requires a trip generation dataset describing socioeconomics on a TAZ level geography. Therefore, a new procedure for estimating this dataset from the RAZ projections of DRAM/EMPAL was required.

- The travel forecasting process generates travel times for TAZ level geography. Therefore, a new procedure for estimating RAZ level travel times from TAZ travel time tables was required.

7.2 Program Organization and Implementation

The application consists of a variety of AMLs,[2] unix scripts, EMME/2 macros, and two special c-programs. The AML which drives the entire application is called ***tlume.aml*** and has been installed on the machine ***magtpo*** in the directory ***/drive2/deapp***.

It is helpful to view the application as consisting of two major modules, one addressing the DRAM/EMPAL side and the other addressing the EMME/2 side of the feedback loop mechanism.

- DRAM/EMPAL Side: ***tlume.aml*** is the driver program for the application. It provides users with an interactive graphic user interface (GUI) of menus and forms necessary to run DRAM/EMPAL. In addition, it calls another driver program, ***runall***, to operate the EMME/2 side of the application.

- EMME/2 Side: ***runall*** is a unix script that runs EMME/2 through a set of special programs and EMME/2 macros. Once completed, written results are left in files which can be accessed once again by ***tlume.aml***.

One issue that had to be addressed concerned the fact that MAG's license for EMME/2 resides on a different machine (***redmtn***) than does MAG's license for both ARC/INFO and DRAM/EMPAL (***magtpo***).[3] There is no common machine on which the EMME/2, ARC/ INFO, and DRAM/EMPAL licenses resided. These licenses are all cpu-locked, so it is necessary to

---

[1] Putnam has long argued that DRAM/EMPAL forecasts of urban growth should be driven by PM peak period travel times.

[2] AML is ARC/INFO's macro language.

[3] A second ARC/INFO license has been subsequently located on the MAG machine redmtn. The DRAM/EMPAL license has subsequently been moved to paradise.

login on the correct machine to access the appropriate licenses, all of which are exercised at various points in the application.

All MAG machines are networked, however, so the application was implemented as follows:

- The main application ***tlume.aml*** runs on the machine ***paradise***, accesses the DRAM/EMPAL license there and ARC/INFO on ***redmtn***.

- The driver for the EMME/2 module, ***runall***, resides on ***redmtn***. It is accessed from ***tlume.aml*** on ***paradise*** through a unix remote procedure call.

- The two modules communicate with each other by reading/writing files in a common directory which can be seen (with both read and write permissions) from *both* ***redmtn*** and ***paradise***. This common directory is actually located on ***magtpo***, and is known locally there as ***/drive2***. From the machine ***redmtn***, this same directory is known as ***/magtpo***.

Unfortunately, since the same directory path names on the MAG network are not known by the same names from different machines, the discussion is somewhat confusing. From now on this will refer to the location of files in terms of the way they are known from the machine ***redmtn***.

7.2.1 Directory Structure

Figure 7-2 illustrates the directory structure for the application. Note that the DRAM/EMPAL side of the application all exists in the directory ***/magtpo/deapp/tlume*** and the EMME/2 side of the application all exists in the directory ***/magtpo/deapp/ emme2***. Table 7-1 describes the general contents of each of these directories.

The source codes for the macros, scripts, and programs comprising the application are all heavily documented, so they are all available for inspection and maintenance. The programs have all been written using structured programming techniques and are modular to aid in their understandability. These source codes can also be read to identify the input/output requirements for their operation.

7.3 DRAM/EMPAL-Side Implementation

The DRAM/EMPAL side of the application did not involve the development of any new methodology — essentially the process for running and managing DRAM/EMPAL that formerly was done on DOS was moved to ARC/INFO. The implementation mirrors the process that was implemented before; therefore, will not provide additional documentation of it. The new ARC/INFO implementation, however, did involve the design of a number of new menus and graphic interfaces that will be discuss.

The ***tlume.aml*** is the interactive interface designed to let users run the entire DRAM/EMPAL process on the unix workstation. It provides the user with the ability to

Figure 7-2
Directory Structure for the Application

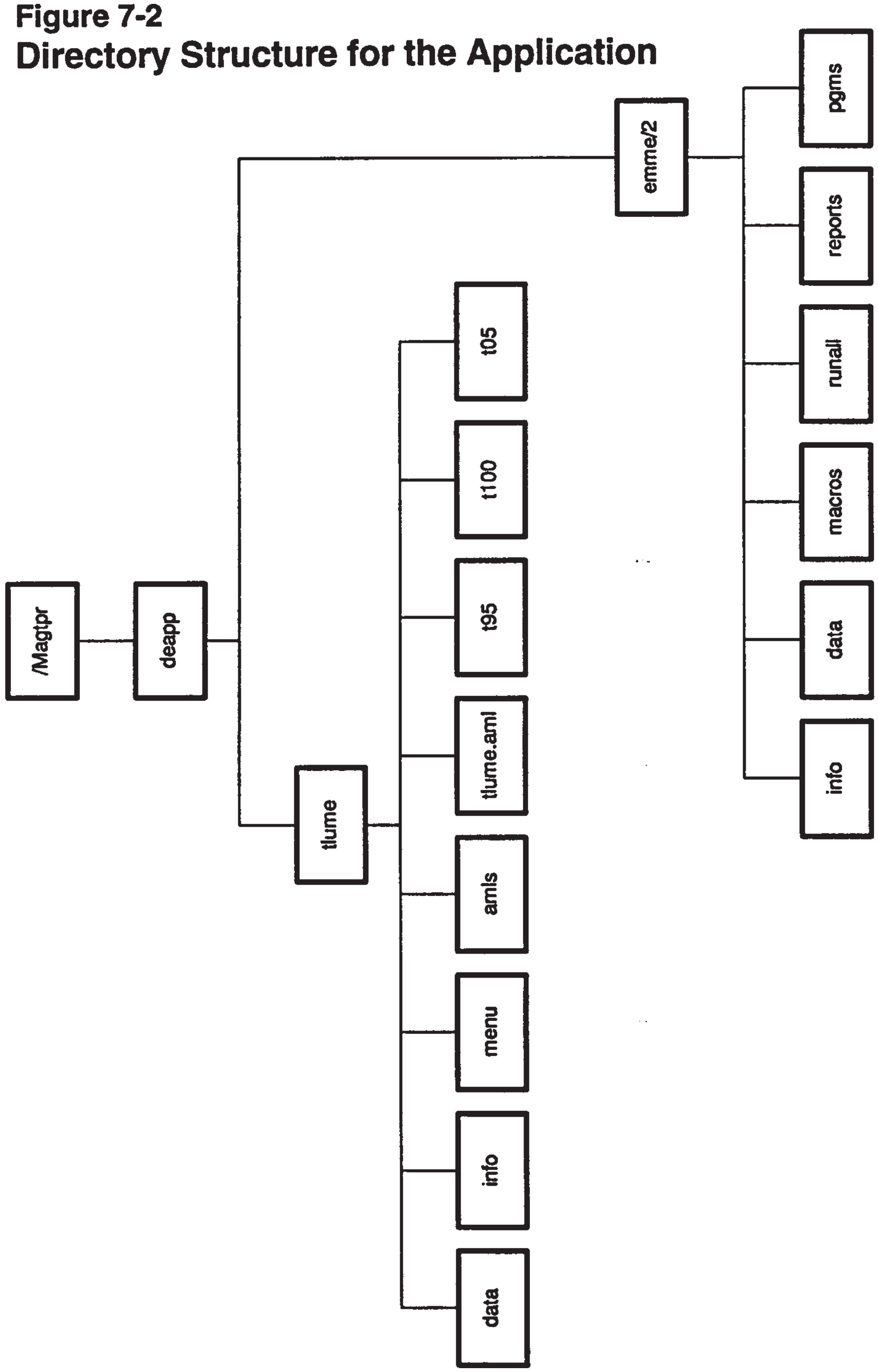


Table 7-1
General Directory Contents

Side	Directory	File or Subdirectory	Description
DRAM/ EMPAL	/deapp	dramempal	All of the licensed DRAM/EMPAL software is installed here.
	/deapp/tlume	tlume.aml	Main driver program for the application
		amls	Supporting amls called by tlume.aml to complete the application reside here
		menus	Menus generated by tlume.aml to provide the user interface reside here
		info	The INFO subdirectory containing INFO files used by the application
		95, 00, 05	Subdirectories containing datasets relevant to the forecast years in question; these are created by the user as the application progresses
EMME/2	/deapp/emme2	runall	Main driver program for EMME/2 (a unix script)
		pgms	Special c-programs, notably tazdata.c and ttimes2.c which are needed for data conversions to support the application all reside here
		macros	Contains macros peculiar to EMME/2 model operation in support of DRAM/EMPAL. Usually regular MAG EMME/2 macros in /mag/emme2/macros are used unless the application required revisions
		data	Miscellaneous ASCII datasets common to all application runs. Contains RAZ lookup tables, etc.
		reports	All reports generated during an EMME/2 run are left here.
		info	INFO directory contains INFO database versions of the ASCII files contained in the data subdirectory reside here

create base year and forecast year runs of DRAM/EMPAL and to complete runs of the travel forecast model needed for DRAM/EMPAL forecasts. The process that was implemented in DOS was replicated in the unix and INFO environment. It was also necessary to recompile a number of FORTRAN programs for converting DRAM/EMPAL datasets between ASCII and binary formats (***dataprep, dascii,*** and ***eascii***) for the unix workstation.

While both DRAM/EMPAL and the travel demand forecasting process have been set up to run entirely from the ***tlume.aml*** application, the user still has the flexibility to modify the input files (e.g. the DRAM/EMPAL card files) using a unix text editor. It is also possible to edit the input datasets with a text editor (when they are in ASCII format) or INFO.

Figure 7-3 shows the main menu of the TLUME application; Table 7-2 explains the functions of the buttons in the main menu. Secondary menus appear in Figure 7-4.

Table 7-2
TLUME.AML Commands

Button	Function
Log Directories	Let the user name the location of model run results.
Create Base	Creates a Base Year Subdirectory.
Initialize Base	Copies Base Year files/databases to the Base Year Subdirectory and runs the dataprep programs (see Figure 7-4a).
Calibrate EMPAL	Checks for necessary files and runs EMPAL calibration (see Figure 7-4b).
Calibrate DRAM	Checks for necessary files and runs DRAM calibration.
EMPAL K-Factors	Checks for necessary files and runs EMPAL K-fac (see Figure 7-4c).
DRAM K-Factors	Checks for necessary files and runs DRAM K-fac.
Delete Base	Deletes a Base Year Subdirectory.
Batch Forecast Run	Runs the dram/empal chain of programs to produce a forecast.
Inner Loop Iterations	Sets the number of iterations for the DRAM and EMPAL chain of programs.
Save Loop Results	The results for each DRAM and EMPAL iteration will be saved if the box is checked.
Create Forecast	Creates a Forecast Year Subdirectory.
Initialize Forecast	Copies Forecast Year files/databases to the Forecast Year Subdirectory and runs the dataprep programs (see Figure 7-4d).
EMPAL Card	Check for the existence of necessary EMPAL card files and other databases (see Figure 7-4e).
DRAM Card	Check for the existence of necessary DRAM card files and other databases.
EMPAL	Generates an employment forecast for a forecast year (see Figure 7-4f).
DRAM	Generates a household forecast for a forecast year.
Delete Forecast	Deletes a Forecast Year Subdirectory.
Run EMME2	Copies necessary files and runs the EMME2 macro runall.
Quit	Ends the tlume application session.

Figure 7-3

The *tlume.aml* Main menu

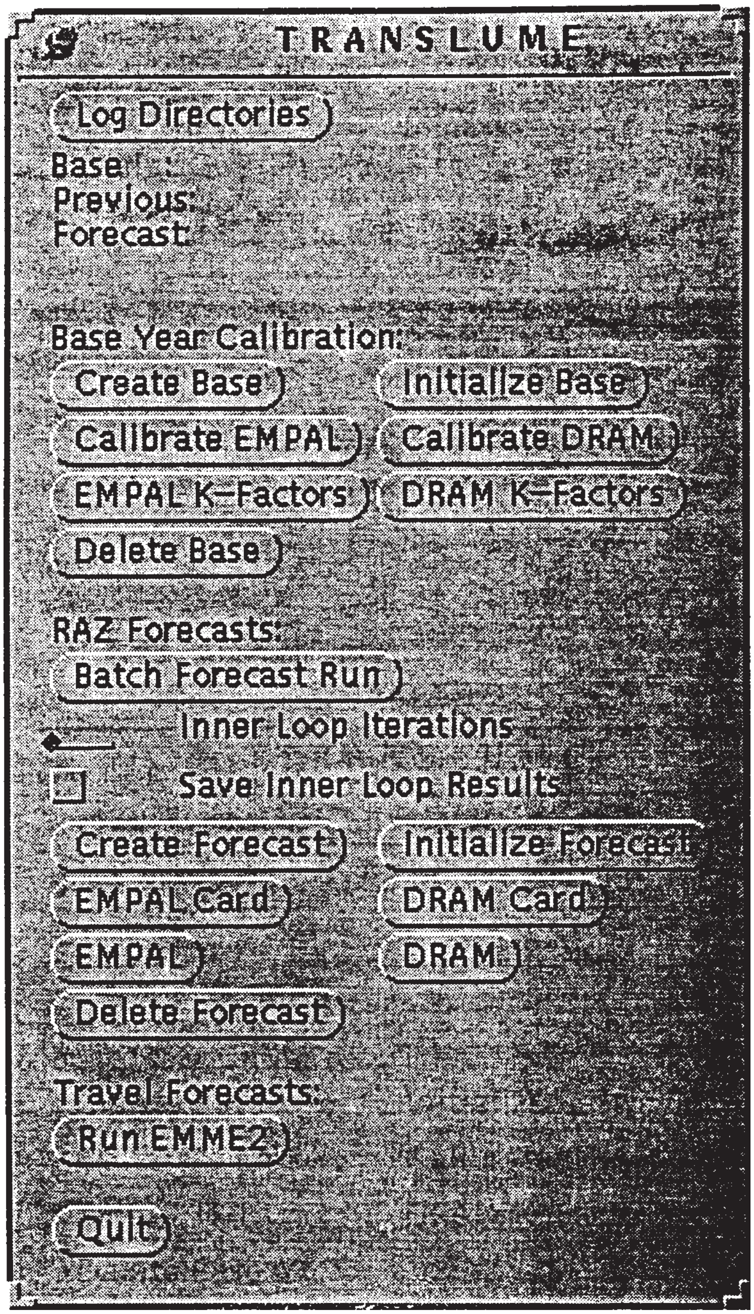


Figure 7-4
Secondary tlume.aml Menus

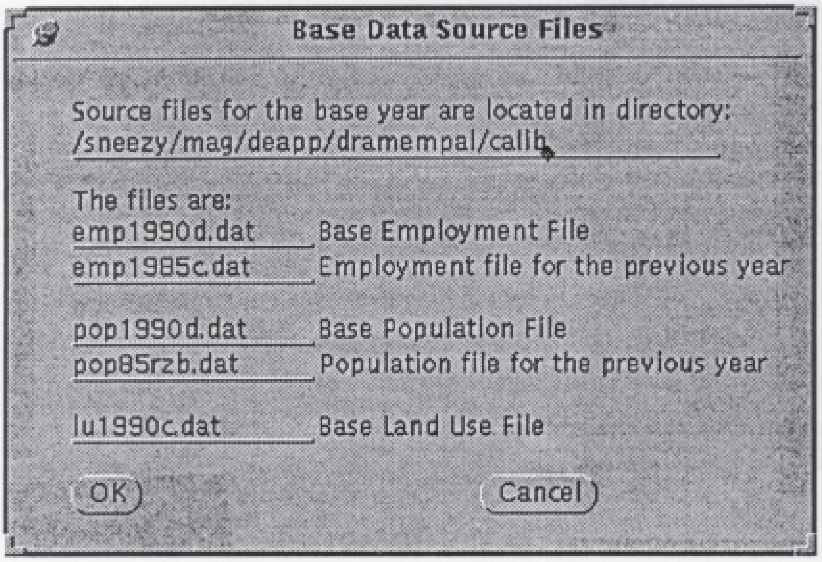


Figure 7-4b
Secondary tlume.aml Menus

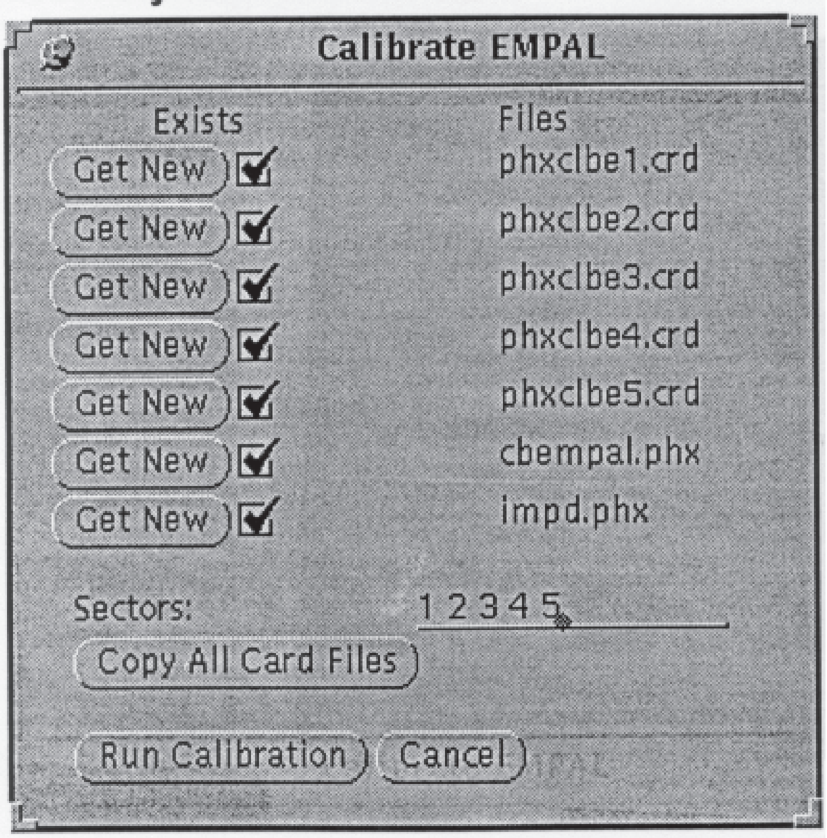


Figure 7-4c
Secondary tlume.aml Menus

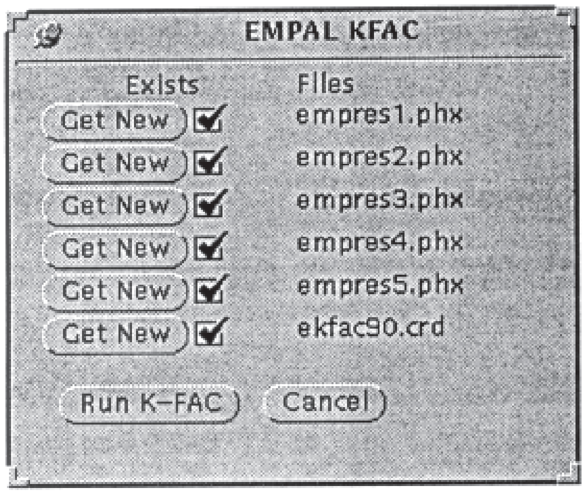


Figure 7-4d
Secondary tlume.aml Menus

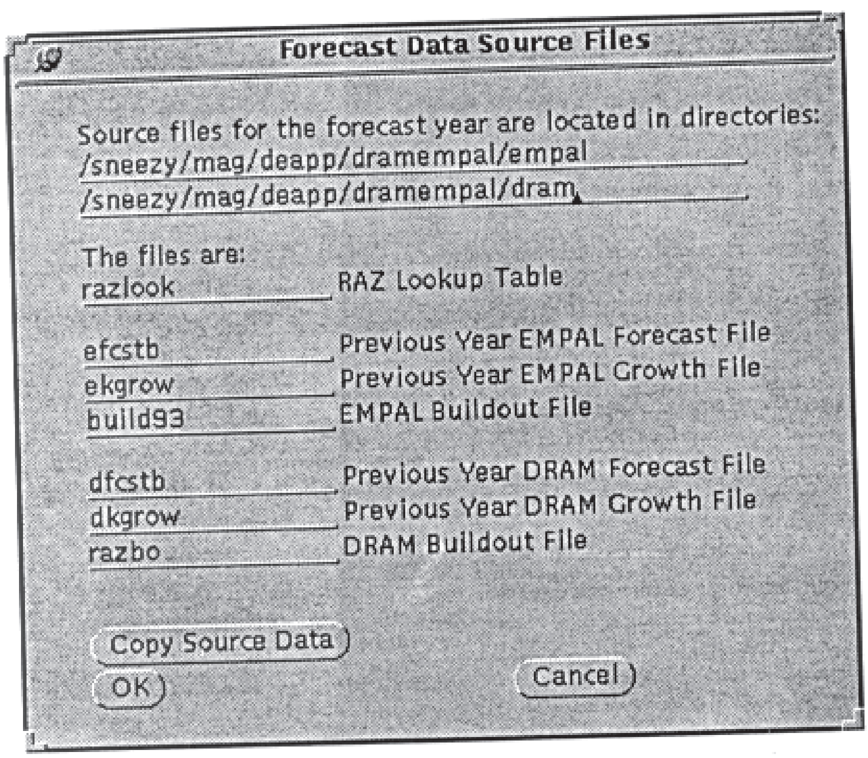


Figure 7-4e

Secondary tlume.aml Menus

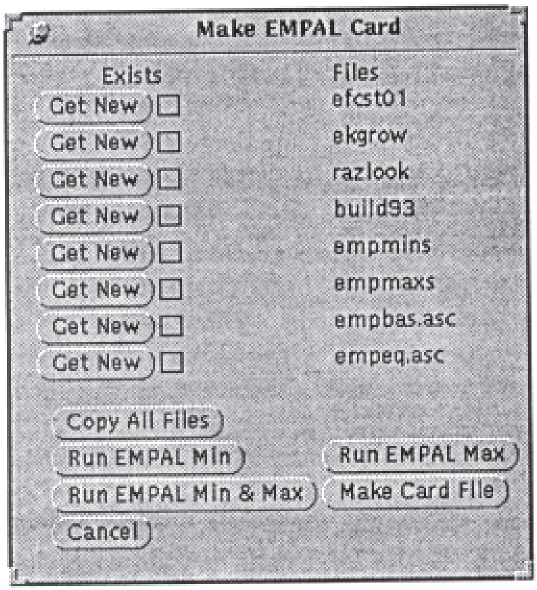
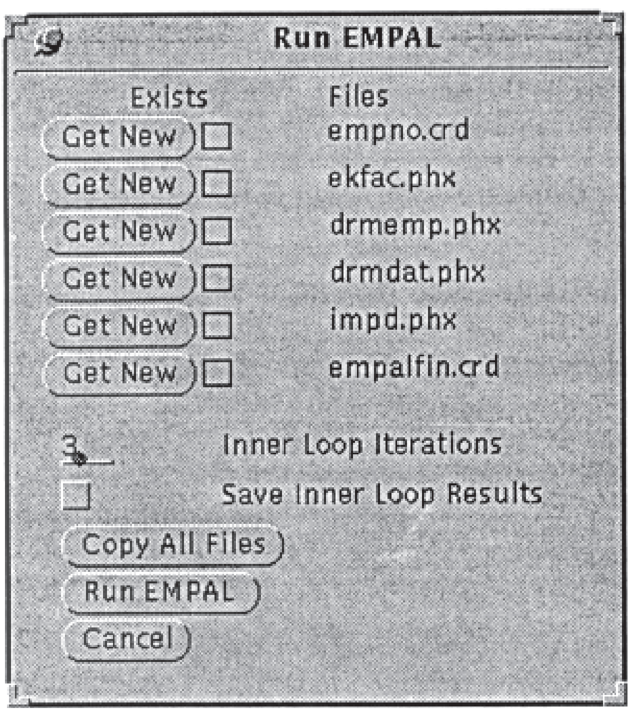


Figure 7-4f

Secondary tlume.aml Menus



TLUME.AML Menus

- Figure 7-4a shows the *Base Year Initialization* menu. The input line on the menu shows the location (at the top of the menu) and names of the base year datasets. When the **OK** button is selected the files are copied to the base year directory and the ***dataprep*** and ***ttimes2*** programs are run.

- Figure 7-4b shows the *Calibrate EMPAL* menu. The menu shows the names of the card files and other files needed to calibrate EMPAL. If the file exists in the base year subdirectory, the checkboxes are marked appropriately. If the boxes are not marked, the **Get New** button can be selected and the user will be prompted for the location and name of the card or other file. The Sectors input line is used to indicate which employment sectors will be calibrated. The user need only calibrate those sectors which have not been calibrated so far (the desired sectors should be separated by a space). A similar menu for the DRAM side will appear when the **Calibrate DRAM** button is selected in the main menu.

- Figure 7-4c shows the *EMPAL K-factor* menu. The menu shows the names of the files necessary for running the K-fac program. If the file exists in the base year subdirectory, the checkboxes are marked. If the box is not marked, the **Get New** button can be selected and the user will be prompted for the location and name of the file. The **Run K-FAC** button runs the ***k-fac*** program and keeps the output files in order. A similar menu for the DRAM side will appear when the **DRAM K-Factors** button is selected in the main menu.

- Figure 7-4d shows the *Forecast Data Source Files* menu which appears when the **Initialize Forecast** button is selected in the main menu. When this button is selected, the forecast subdirectory is cleaned out and the user has the opportunity to copy over new files to the subdirectory. The names of various preexisting files that are needed for the forecast are found in the input fields of the menu.

- Figure 7-4e shows the *Make EMPAL Card* menu. This menu shows the names of the files that are needed to create the final EMPAL card file. The first four files (***efcst01, ekgrow, razlook, build93***) are needed to find the minimums and maximums. The next two files (***empmins*** and ***empmaxs***) are created when the **Run EMPAL Min, Run EMPAL Max** or **Run EMPAL Min and Max** buttons are selected. The next two files (***empbas.asc*** and ***empeq.asc***) are needed with the mins and maxs files to run the **Make Card File** button. A similar menu for the DRAM side will appear when the **DRAM Card** button is selected in the main menu.

- Figure 7-4f shows the *Run EMPAL* menu. The menu shows the files necessary to do an EMPAL run. If the user has stepped through the buttons from creating the forecast directory to creating the card files, all of the check boxes should be marked. The user also has the opportunity to change the number of inner loop iterations and to save the inner loop results in this menu. The user can also copy the necessary files over with the **Copy All Files** button.

7.4 EMME/2-Side Implementation

The ***runall*** script is an executable unix-script file that is responsible for running the entire EMME/2 travel demand forecasting process, including the conversion of datasets between forms that can be handled by EMME/2 and forms that can be handled by DRAM/EMPAL. The approach was to rely on EMME/2 macros used in regular modeling applications to the extent possible (and, in fact, those macros are called directly); special versions of these macros when the procedures specifically oriented to DRAM/EMPAL feedback loop requirements deviate from the norm.

An overview of program flow is provided in Figure 7-5. It involves the following steps:

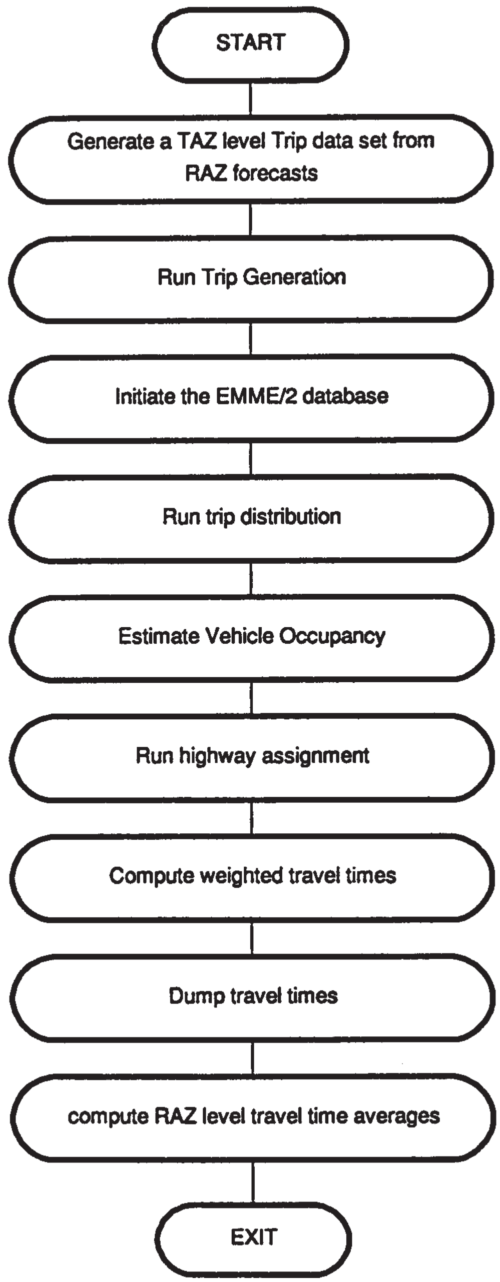
- ***tlume.aml*** writes several source files to the ***/magtpo/deapp/emme2*** directory that will be needed by the application. These are covered below.
- ***tlume.aml*** executes a remote procedure to ***runall***, which sets the EMME/2 forecasting process in motion.
- ***runall***calls ***tazdata.c***,[4] which converts DRAM/EMPAL forecasts of households and employment to a new trip generation dataset needed by the travel forecasting process. Details on this follow.
- ***runall*** runs the entire EMME/2 forecasting chain from trip generation through highway assignment, but for the PM peak period only and omitting mode choice.
- ***runall***averages travel volumes with any previous iterations and recomputes link travel times and issues a TAZ level travel time matrix (and a trip table used to weight travel times).
- ***runall***calls ***ttimes2.c*** to calculate average travel times for RAZ level geography. The binary version of the results are written to the file ***impd.phx*** which can then be read by ***tlume.aml***
- ***runall*** exits and control reverts to ***tlume.aml.***

7.4.1 EMME/2 Databank Setup

For this application, the EMME/2 databank does not have to be as large as it is to support regular MAG travel forecasting runs — the main difference is that the databank need only

[4] This actually is the source code. The executable file is known as **tazdata**, but this will be continued to be referred to the source-code name to make it clear that referring to a custom-written c-program.

Figure 7-5
The EMME/2 Forecasting Process



support two different scenarios. The parameters defining the databank that is needed are shown in Table 7-3. The macro ***newdebank.mac*** that have been provided will establish the databank correctly.

Table 7-3
EMME/2 Databank Setup

Number of Scenarios	2
Number of Centroids	1,288
Number of Nodes	12,800
Number of Links	32,000
Number of Turns	15,000
Number of Transit Vehicles	1
Number of Transit Lines	1
Number of Line Segments	1
Number of Matrices:	
ms	75
mo	75
md	75
mf	99
Functions per Class	99
Operators per Class	2,000
Log Book Size	100
Demarcation set size	100
Extra Attribute Storage	32,001
Node Labels	No
User Data on Transit Line Segments	No
Project Title	DRAM/EMPAL
Terminal Type	2 (Emtool)
Printer Type	4 (HP Laserjet)
Plot file Type	1

7.4.2 EMME/2 Source Files

The ***runall*** script and the EMME/2 batch process requires ***tlume.aml*** to write several files to the ***/magtpo/deapp/emme2*** directory. These files are shown in Table 7-4.

Table 7-4
Source Files Needed by the EMME/2 Driver *RUNALL*

File	Description
dfcst.dat	ASCII file of households by income class generated by the last iteration of DRAM/EMPAL. Households are listed for each RAZ by income class.
efcst.dat	ASCII file of employment by employment type generated by the last iteration of DRAM/EMPAL. Employment classes are listed for each RAZ with employment types appearing in the following order: other, government, retail, office, and industrial.
year.dat	A single entry file showing the forecast year. Used by the tazdata.c program to interpolate growth in special populations from control totals appearing in the file regcontrol.dat.
network.dat	A copy of the network appropriate to the forecast year; copied from the directory /mag/modeldat/networks.
d231.in	Stock d231.in file used by the modeling process.

Once this is accomplished, the ***runall*** script will run. In addition to various report files which are created during the process, the output files are:

- impd.phx: RAZ-to-RAZ travel times in binary format suitable for input to the DRAM/EMPAL program.[5]
- impd.dat: An ASCII version of the same table.

In addition, the process relies on a number of resource files which reside in the ***/magtpo/deapp/emme2/data*** subdirectory. These were shown in Table 7-5.

7.4.3 Special EMME/2 Feedback Macros

As stated, the EMME/2 model chain driven by ***runall*** calls regular EMME/2 macros in the ***/mag/emme2/macros*** subdirectory to the extent possible. There are variations in the implementation to serve the purposes of the DRAM/EMPAL feedback loop, and some macros were revised and/or augmented. These are all located in the ***/magtpo/deapp/emme2/macros*** subdirectory and are described in Table 7-6.

The key methodological differences which are reflected in these macros are as follows:

[5] An unexpected idiosyncracy in the binary format for this file has been discovered. It appears not to conform to regular unix binary files, but has a format reminiscent of older FORTRAN file structure (which was the source code language for DRAM/EMPAL). Check the program code for ttimes2.c for details.

Table 7-5
Resource Files Needed by the EMME/2 Driver *RUNALL*

File	Description
basetaz.dat	A base year TAZ level trip generation data set; taz1993a.txt was used. The file is used as a basis for determining growth factors from the DRAM/EMPAL run applicable to a number of variables (e.g., population).
razlook.dat	A file describing RAZs, including their public and internal identification number, their x and y coordinates, and a marker to flag RAZs which are external to the modeling area.
fullook.dat	A file describing the TAZ-to-RAZ lookup table.
regcntl.dat	A file of countywide control totals for special populations (seasonal, transient, and group quarters).
xneighbor.dat	A file that describes, for each RAZ external to the modeling area, the closest neighbor that resides internal to the modeling area.

Table 7-6
Special EMME/2 Macros Developed for the Feedback Loop Mechanism

Macro	Description
newdebank.mac	This macro establishes a new EMME/2 databank to support DRAM/EMPAL feedback loop forecasting.
setup.mac	This macro sets up a new EMME/2 databank by installing appropriate units of measure and other duties.
init.mac	This macro precedes every EMME/2 run; it essentially deletes scenario 2 so that it may be used again.
hskims.mac	This macro runs highway skims.
dist.mac	This macro runs the original method for trip distribution (weighted generalized cost skims are now being run by MAG). The macro is merely a copy of the original.
vehocc.mac	This macro computes total PM peak period vehicle trips from person trips carried through trip distribution, thereby replacing the mode choice process which is normally run at MAG.
hassign<*i*>.mac	This macro runs highway assignment for the PM peak period. It then computes average highway volumes for iteration *I* based on the current and previous results, storing the weighted volume as link attribute UL2 and as the extra attribute @pmvol. New volume-delay functions are read (that key off of UL2) to compute new travel times.
ttimes.mac	This macro dumps the travel time matrix (mf33) and the vehicle trip matrix (mf55) to ASCII tables for processing by the ttimes2.c program.

- As noted before, the stock EMME/2 databank which MAG normally establishes for a typical modeling run is larger than necessary to support the DRAM/EMPAL feedback loop mechanism. Our version only requires a databank with 2 scenarios and eliminates the need for storage for transit information.

- A vehicle occupancy macro to directly convert daily person trips to PM peak period vehicle trips was developed, thereby obviating the need for mode choice analysis. Peak period factors were based on the original MAG implementation; vehicle occupancy factors by each trip purpose were based on end-results reported from full MAG mode choice runs.

- A new highway assignment macro was developed which implements several changes:

    - It runs only PM peak period highway assignment (AM peak and daily runs are not required).

    - It implements the convex function weighting scheme discussed earlier. New volumes are stored in UL2 and saved to the extra link attribute @pmvol for use on the next iteration. New link travel times are computed based on the volume in UL2 with a new set of volume delay functions (which are essentially the same as the originals, except they reference UL2 rather than volau).

    - In order to force the generation of a travel time matrix (instead of a generalized cost impedance matrix), this macro makes one final run of a null scalar trip table matrix with volume delay functions keying off of UL3, which stores congested travel times from highway assignment.

- A new macro was developed to dump the resulting trip table and travel time matrix to ASCII files for conversion to RAZ level geography.

7.4.4 TAZ Socioeconomic Conversion Program: tazdata.c

Intermediate DRAM/EMPAL forecasts are written to the ***/magtpo/deapp/emme2*** directory as the ASCII text files ***dfcst.dat*** (for DRAM) and ***efcst.dat*** (for EMPAL). These must be disaggregated from RAZ level geography to TAZ level geography. In addition, the trip generation data set used by the travel forecasting model contains a number of other variables that do not come from DRAM/EMPAL. These must be generated as well, resulting in the dataset ***tazdata.dat***. This is the job of the program ***tazdata.c***.

The trip generation dataset is shown in Table 7-7. Variables identified in the table which come directly from DRAM/EMPAL (although for aggregate RAZs) include (1) households and (2) each of the five employment classes. Other variables, such as household population, special populations, post high school enrollments, etc., need to be derived.

Table 7-7
The Trip Generation Dataset

Field	From DRAM/ EMPAL?	Format	Description
YEAR	n/a	I4	Forecast Year
TAZID	n/a	I5	TAZ number
DISTRICT	n/a	I4	District number
MPA	n/a	2C	Metropolitan Planning Area Identifier
HHPOP		I6	Resident Population in Households
GQPOP		I5	Population in General Quarters
TRANPOP		I5	Transient Population
SEASPOP		I5	Seasonal Population
HH	yes	I6	Resident Households
GQHH		I5	General Quarters Households
TRANHH		I5	Transient Population Households
SEASHH		I5	Seasonal Population Households
OTHEMP	yes	I6	Other Employment
GOVEMP	yes	I6	Government (or Public) Employment
RETEMP	yes	I6	Retail Employment
OFFEMP	yes	I6	Office Employment
INDEMP	yes	I6	Industrial Employment
HHINC		I7	Median Household Income (1989 $)
DEVRESAREA	*	I6	Developed Residential Area (square miles * 100)
UNDRESAREA	*	I6	Undeveloped Residential Area (square miles * 100)
DEVEMPAREA	*	I6	Developed Employment Area (square miles * 100)
UNDEVAREA	*	I6	Undeveloped Area (square miles * 100)
AREA	*	I6	Total TAZ Area (square miles * 100)
VEHICLES	*	I6	Total Vehicles in TAZ
UPPERENROL		I6	Post High School Enrollment by Place of Residence
RETIRE	n/a	I2	Retirement Flag
ASU_DIST	n/a	I4	Distan/ae to ASU (miles * 10)
DAILYPCOST	n/a	I4	Average Daily Parking Cost in TAZ (cents)
HOURLYPCOST	n/a	I4	Average Hourly Parking Cost in TAZ (cents)
ENPLANEMENTS		I6	Average daily enplanements at Sky Harbor
TERMTIME	n/a	I2	Terminal Time (minutes)

n/a: Not applicable: coding remains as in the base year
* Not used.

Base Year Variables

The tazdata.c program reads a base year trip generation dataset from the ***/magtpo/deapp/emme2/data*** subdirectory. The ***taz1993a.dat*** dataset to represent the base year was used, but any one will do. This base year trip generation dataset is used as a source for a number of variables for the forecast year which should not be expected to change. These would include the TAZID number, the DISTRICT, the MPA, the RETIRE flag, the ASU_DIST variable, and the TERMTIME. In addition, tazdata.c leaves DAILYPCOST and HOURLYPCOST as they are in the base dataset.

Disaggregating DRAM/EMPAL Variables

Projected values for variables for which DRAM/EMPAL is a source are prorated to each TAZ on the basis of the TAZs share in the base year. This applies to households and each of the employment variables. The ***tazdata.c*** program reads the TAZ-RAZ lookup table (***fullook.dat***) and therefore knows the TAZ membership of each RAZ.

Using households as an example, the number of households for time t in TAZ I is proportional to the growth which has been exhibited for RAZ m, assuming that TAZ I is in RAZ m.

$$HH_i^t = HH_i^{t-1} \frac{HH_m^t}{HH_m^{t-1}}$$

Generating Derived Variables

The method used to generate estimates of derived variables is as follows:

- Household Population

  Household population is factored upwards in proportion to growth in households in the TAZ. This is, whatever persons per household rate was exhibited in the TAZ in the base year applies equally to the forecast year.

- Special Populations

  A file (***regcntl.dat***) of control totals by forecast year for special population groups (seasonal, transients, and group quarters) is available to the program in the ***/magtpo/deapp/emme2/data*** subdirectory. Countywide totals are computed for the forecast year (by interpolation) from this control file in order to establish growth factors in comparison with the base year. These growth factors are then applied across the base year dataset to create estimates for the forecast year.

- Enplanements and Post High School Enrollment

  An overall growth factor in Maricopa County households for the forecast year is computed in comparison with households in the base year. This growth factor is then applied to these two variables.

In the current version, median household income for each TAZ is left as it was during the base year. This will preserve the breakdown of households by quintiles throughout the County.

Data Checks

The results of the ***tazdata.c*** program were closely checked for 1995 DRAM/EMPAL runs in comparison with the ***taz1995a.dat*** dataset. Overall totals for the county, as well as for individual RAZs, were found to be accurate.

7.4.5 Generating Travel Times With ttimes2.c

Once travel times are generated on TAZ level geography by the EMME/2 forecasting model chain, they must be converted to RAZ level geography so they can be input to DRAM/EMPAL. This is the job of a special c-program ***ttimes2.c***.[6] This program reads two batchout files from EMME/2 (***ttrips.dat*** and ***ttimes.dat*** generated by the ***ttimes.mac*** EMME/2 macro) along with a number of resource files (***razlook.dat, fullook.dat***, and ***xneighbor.dat***) contained in the ***/magtpo/deapp/emme2/data*** directory and calculates inter-RAZ travel times from them. The results are posted to the files ***impd.phx*** and ***impd.dat*** (the ASCII version) and summarized in the report ***impd.rpt***. The ***tlume.aml*** AML can then read the new travel times and set out on another iteration of DRAM/EMPAL.

Initially, the ***ttimes*** program was developed to mirror the procedures which were originally used by MAG. The current version ***ttimes2*** substantially improves on those procedures, as follows:

- Internal RAZ travel times[7] formerly were estimated by selecting a representative TAZ which resided in the RAZ in question — and there was a file declaring the representative TAZs for each RAZ. Travel times between these TAZs were then used to represent inter-RAZ travel times.

  This has been modified; now inter-RAZ travel times are computed as the average of all inter-TAZ times which are members of the two RAZs, weighted by the trips involved.

- Travel times from RAZs external to the modeling area formerly were based on the airline distance between the centroids of both, traveling at an average speed of 35 mph. Also, a supplemental set of distances was added to make some RAZs appear farther apart then they really were. This technique was apparently developed in order to improve the results of DRAM/EMPAL, however its precise origin is not documented.

---

[6] This is a second version of the program. After investigating the treatment of externals in the first version, ttimes.c, this new version was implemented.

[7] Recall that while most RAZs reside internally to the modeling area, there are 18 of them that are external to the modeling area.

This has also been modified. Now a file defining the internal RAZ neighbor for each external RAZ is read; the travel times between RAZs consists of the average airline distance time from the external to the internal neighbor plus the normal travel times between the two internal RAZ neighbors. This method assures that any congestion occurring in the metropolitan area affects external RAZs as well.

Handling Internal Trips

As mentioned, travel times between internal RAZs are computed as the weighted average among all the TAZs which are members of either RAZ — weighted by the number of trips which experience those times.

In other words, the travel time TT_{mn} between two RAZs m and n is:

$$TT_{mn} = \frac{\sum_{ij} TT_{ij} Vol_{ij}}{\sum_{ij} Vol_{ij}}$$

where TT_{ij} is the travel time between TAZs I and j and Vol_{ij} is the travel volume between I and j, and where TAZ I is a member of RAZ m and TAZ j is a member of RAZ n.

Handling External-Internal Trips

Each external RAZ is associated with an internal RAZ, its nearest neighbor which resides inside the modeling area. The travel time between then can be computed assuming an average speed of 35 mph over the straightline airline distance between them. Therefore, the travel time TT_{mn} from an external RAZ m which has an internal neighbor k to an internal RAZ n is:

$$TT_{mn} = 60\frac{D_{mk}}{S} + TT_{kn}$$

where D_{mk} is the airline distance between RAZ m and RAZ k and S the assumed speed of 35 mph.

Handling External Through Trips

Since the RAZ area includes the entire County which is much larger then the modeling area, in some cases, external through trips would be expected to traverse the modeling area; in some cases, they would not. Therefore, each external RAZ is also coded with a group number — external RAZs which belong to the same group would not be expected to traverse the modeling area. Their travel times are predicated only on airline distance, as in:

$$TT_{mn} = 60\frac{D_{mn}}{S}$$

Travel times between two external RAZs that belong to different group numbers, however, can be expected to travel through the modeling area. Therefore, the travel time between two external RAZs m and n which respectively have two internal neighbors k and l is:

$$TT_{mn} = 60\frac{D_{mk}}{S} + TT_{kl} + 60\frac{D_{ln}}{s}$$

Data Checks

The ***ttimes2.c*** program generates a summary report of results, which appear to be reasonable. Average speeds in the region are on the order of 25 mph to 35 mph. The results of this new methodology was also explored in comparison to the former methodology for the year 1995, with the following results:

- The method of weighting internal TAZ travel times on the basis of travel volume did not offer results much different than a straight arithmetic average would provide.
- The method of basing travel times on all TAZ members of a RAZ did not substantially change the results, compared with using representative TAZs.
- Differences between times now produced and those on which DRAM/EMPAL has been based historically are more likely to be the result of changes in MAG's travel forecasting procedures and the use of EMME/2, rather than its previous DRAM/EMPAL input forecasting techniques.
- Differences between RAZ to RAZ travel times are not more than 5-10 minutes, and the new results look more reasonable.

8
The Subarea Allocation Model

Two geographies are generally used in forecasting activities at MAG. *Regional analysis zones* (RAZ) support the land use forecasting process operated with the software package DRAM/EMPAL. Associated with DRAM/EMPAL is a database describing (1) household projections by income class (quintiles) and (2) employment projections by employment type.

The *traffic analysis zone* (TAZ) system is the *other* geography in use at MAG. The TAZ geography is more detailed than the RAZ geography, and is used by the travel demand forecasting process operated with the software package EMME/2. The TAZ database associated with the travel demand models includes the household and employment variables generated by DRAM/EMPAL, and many additional variables.

Both of these zone systems were addressed in Chapter 3.

Converting land use projections generated by DRAM/EMPAL on a RAZ geography to comparable projections needed by trip generation on a more detailed TAZ geography presents MAG with significant problem of *disaggregation*. How do you allocate 1,000 households in a RAZ to the 10 constituent TAZs that comprise it?

In addition to this, though, MAG is often requested to generate representations of forecast land use data sets on other geographies; for example, municipal or water district boundaries. These geographies rarely conform to either TAZ or RAZ zone boundaries and therefore, meeting these requests can present considerable computational difficulties, including zone splitting and additional disaggregations.

The solution envisioned by MAG in this project is to create a new zone system called a *minimal analysis zone* (MAZ) system which would serve as a base geography on which all land use forecast information would be carried and which would be *upwardly compatible* with *any* other geography for which land use summaries were needed. This would reduce *all* geographic data problems to a simple data aggregation operation.[1]

As discussed in Section 4, the approach that was taken is to build on ARC/INFO's GRID capabilities to deal with raster geographies. This is to say that the MAZ is not a traditional polygon zone system at all. Instead, the objective is to build a *rasterized* geography of grid cells (like pixels on a computer screen) covering Maricopa County.

The **Subarea Allocation Model** (SAM) described in this section provides MAG with a method for disaggregating land use and socioeconomic information to the MAZ structure. It also provides MAG with a capability to aggregate MAZ land use data to any other geography of interest, to include the RAZ and TAZ structures, but also any others including future geographies that have not been implemented. The SAM, based on raster geography, is a powerful concept — it not only responds to MAG's immediate needs to convert DRAM/EMPAL forecasts to TAZ trip generation databases, but also allows MAG to automatically convert these datasets to new geographies (such as would be required to support windowed traffic modeling undertaken by communities in the region).

The SAM represents a significant innovation in the use of GIS to address the significant computational difficulties associated with data disaggregation; a problem which is common to MPOs across the country.

8.1 Objectives

The objective of this task is to develop the Subarea Allocation Model, including the ARC/INFO GIS application and its associated methodologies and databases. The functionality required of the application includes:

- The ability to generate an ARC/INFO grid database (the MAZ) for any polygon or point coverage[2], including migrating any set of attributes associated with the original coverage to the grid.[3]

---

[1] In other words, a data summary could be created by merely adding up the variable for its constituent parts.

[2] The existing land use and RAZ covers are examples of polygon coverages which will need to be gridded. The mobile home/RV park and motel/hotel/resort coverages are examples of point coverages which will need to be gridded.

[3] Note that cover values representing quantities, such as number of households, will have to be converted to cell-based values representing densities, such as households per cell. Cover values representing identifiers or rates, such as RAZ numbers or occupancy rates, should be left alone.

- The ability to aggregate data associated with any gridded geography to any polygon structure. Household quantities associated with grid cells, for example, would be accumulated to form a TAZ dataset.[4]

- The implementation of a methodology for assigning RAZ growth in households and employment from a DRAM/EMPAL run to individual grid cells. This methodology should reflect (1) eligible land to accommodate growth, (2) appropriate general plan designations, (3) locational preferences associated with various residential and employment land uses, (4) knowledge of future projects already approved by municipalities but not yet built, and (5) appropriate land use absorption rates associated with growth.

- The implementation of a similar methodology for assigning countywide growth in special population groups (e.g., mobile home parks, motels, etc.) to individual grid cells.

- The ability to generate other *derived* TAZ variables from those already contained in the grid data structure. These would include variables such as population (from households), and median income (from household income classifications), for example.

- The ability to regenerate new point and polygon coverages from their grid representations. This is required, for example, to establish new *then-existing land use* covers for another series of forecasts.

8.2 General Approach to the SAM

With *no other* additional information, disaggregation of data from one geography to a more detailed geography can not be done in any way other than prorated shares based on area. MAG, however, has developed more information about the distribution and uses of lands within Maricopa County than appear in either the RAZ or TAZ datasets. Some of these coverages include:

- ***exlu***: the existing land use polygon coverage for 1994. This cover was shown earlier in Figure 6-1. Not only do land use codes associated with the cover portray the current distribution of lands by various uses, but they also distinguish areas which are available for development from those which are already built (or ineligible for development). There are roughly 14,000 polygons in this cover, so it provides considerably more detail than the roughly 1,300 TAZs.

- ***gplan***: the general plan polygon coverage assembled from data coming from municipalities within the county. This cover was shown earlier in Figure 6-2.

[4] Similarly, this requires a distinction between cell values that represent quantities from those that represent identifiers or rates.

Growth projected by the DRAM/EMPAL model should be allocated to lands within RAZs in accordance with the policies expressed by the general plan coverage.

The fundamental database concept behind the SAM is the notion of the *then-existing land use cover* applicable to any given forecast year. This cover describes all land uses in the county and is quantified with households, employment, and other relevant variables for each land use polygon.

- Each polygon in the existing land use cover isolates a single homogeneous land use, fully built-out for developed areas and including vacant developable lands. Attributes associated with land use polygons quantify population, employment, and other variables associated with each land use in the county. In its grid form, this cover can be aggregated to whatever polygon zone system is of interest.

- For each forecast year, polygons representing new developments can be added to the existing land use cover, thereby resulting in the creation of a ***new then-existing land use cover*** applicable to the forecast year. This new cover will also be attributed to quantify population, employment, and other relevant variables.

The general plan cover, known project coverages, and coverages which represent locational preferences for various types of land uses, can all be used to select areas for new development. This approach essentially makes the SAM a land use and growth simulation model abiding by general plan designations established by cities, control totals from DRAM/EMPAL, and other factors which logically affect the development of land.

Generalized Model Logic

Figure 8-1 illustrates the general logic of the algorithm for allocating growth, which will now be discussed:

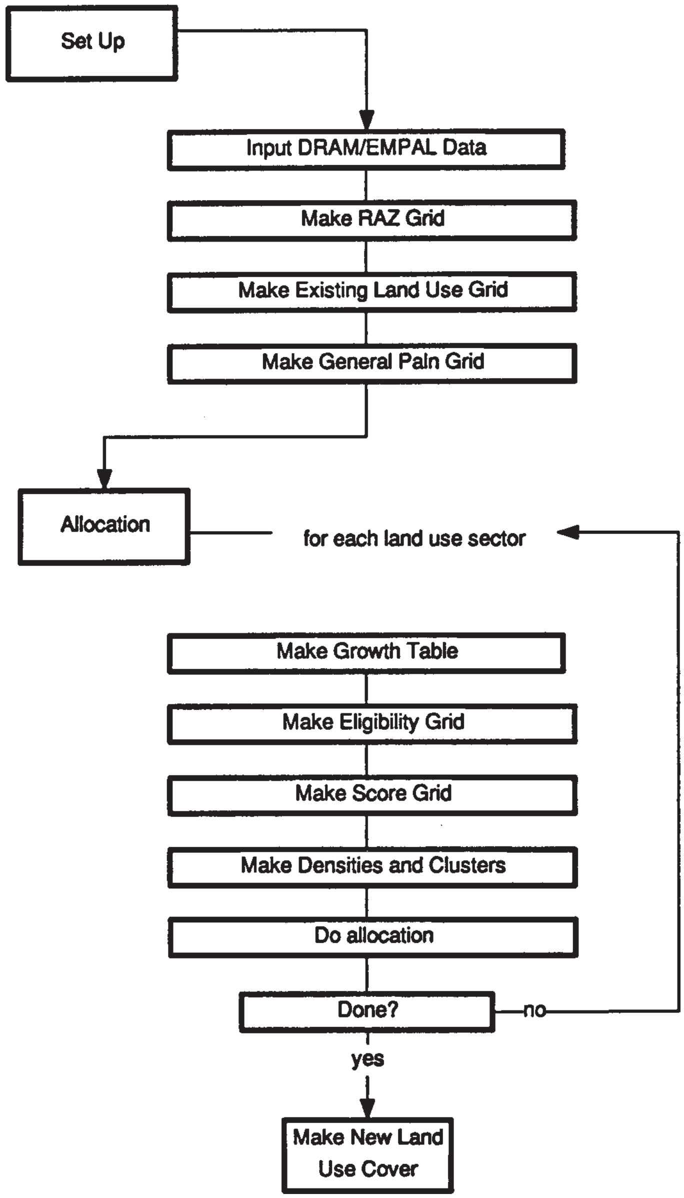
- **Set Up**

  Similar to the implementation of the DRAM/EMPAL feedback mechanism, all of the datasets relevant to a given forecast year reside in a new subdirectory dedicated to that forecast. This subdirectory has to be created. If the user wanted to generate an allocation for the forecast year 2015, for example, they would create a new ARC/INFO workspace (subdirectory) called 2015.

  This directory would then be initialized, involving several operations:

  - The dfcst.asc and efcst.asc ASCII datasets associated with the DRAM/EMPAL forecast would be placed in the directory. They should be imported

Figure 8-1
SAM Flow Chart



into INFO files there and growth files, reflecting changes between the base year and the forecast year, should be created.

- Several coverages should be gridded. These include (1) the RAZ coverage, (2) the base year land use coverage, and (3) the general plan coverage.

- **Running an Allocation**

The application involves the allocation of a number of different land use sectors. For socioeconomic variables that are created in DRAM/EMPAL and allocated from RAZ control totals, these include (1) households and (2) five classes of employment (retail, office, industrial, government, and other). For special population variables that are allocated from countywide control totals, these are (1) motels, (2) mobile homes/RV parks, (3) nursing homes, (4) dormitories, and (5) prisons and jails. No matter what land use sector is being allocated, the basic process is the same, as follows:

- *Make a Growth Table:* A growth table for the land use sector in question is created. For DRAM/EMPAL variables, the growth table contains records for each RAZ, essentially reporting the difference between the DRAM/ EMPAL forecast and the existing land use cover. For special population groups, the table contains only one record, growth in the countywide control total between the base year and the forecast year.

- *Make an Eligibility Grid:* An eligibility grid is created to reflect the eligibility of each cell in the county for absorbing growth of the character being considered. This grid is created by considering two conditions: (1) the availability of undeveloped land as reflected in the existing land use cover and (2) appropriate general plan classification reflected in the general plan cover. If land is both vacant and appropriately designated in the general plan, then it is eligible for development. Of course, lands which are considered eligible for residential growth would be different than those considered eligible for commercial growth by virtue of their general plan designations.

- *Make Scoring Grid:* A scoring grid is created for the land use in question. The scoring grid basically reflects the desirability of cells in the county for absorbing growth of the type in consideration. Each cell is associated with a locational preference score — higher scores essentially reflect land that is more desirable for development.

- *Make Densities:* A density grid is created. The density grid associates a preliminary density with each general plan code. Depending on the amount of growth forecast from DRAM/EMPAL, these densities may be boosted in order to accommodate growth.

- *Make Clusters:* A cluster grid groups adjacent grid cells into blocks and is meant to reflect the fact that development projects are typically larger than a single grid cell. A growth of 1,000 households in a RAZ, for example, might be expected to occur in the form of five 200-unit subdivisions. The notion of clusters implies the fact that if a grid cell is selected for development, then its neighbors will be too because they are part of the same subdivision project.

- *Run Allocation:* The allocation process essentially sorts RAZ cells on the basis of their score. RAZ growth allocation starts with the **highest scoring cluster** of grid cells (as reflected in the *cluster grid* and the *score grid*) which is eligible for development (as reflected in the *eligibility grid*). The grid cell is "developed" at a certain density (as indicated in the general plan and reflected in the *density grid*), absorbing growth, and the allocation then moves to the next highest scoring grid cell. The process continues until all growth is absorbed.

  In cases where there is insufficient land available in the RAZ at the preliminary density levels to accommodate growth, density levels are increased accordingly. In the case of ties, where many grid cells with the same score vie for a finite amount of growth, densities are reduced.

  The process works *exactly* the same, regardless of what land use sector is being allocated or whether it is allocating from RAZ control totals (for DRAM/EMPAL data) or from countywide control totals (for special population groups). Of course, the information in the various grids, such as densities, eligibility criteria, and scores, are different depending on the land use sector in question.

- **Make Existing Land Use Coverage:** All of the grid cells nominated to absorb growth for each of the land use sectors are combined into a single grid. A new growth land use cover is generated from the grid and appended to the old land use cover.

So the driving forces behind the SAM are as follows:

- An existing land use cover reflects developed land. Each polygon appearing in the coverage is populated with all of the forecast variables of interest, whether it be households, employment, motel rooms etc. A series of these coverages are generated, one for each forecast year.

- A general plan cover reflects the eligibility of land for various types of development and the density at which it can be developed.

8.3 SAM Methodologies

An overview of the allocation process was given in the last section. This section will discuss more fully the key elements of the process.

8.3.1 Eligibility of Land to Absorb Growth

The first step in the allocation process requires us to identify land which is available for development, not unlike the manual process that MAG has followed in the past. There are two conditions that must be satisfied for land to be considered for development: (1) it is currently vacant and (2) it is appropriate by virtue of its general plan designation.

On inspection, the existing land use cover maintained by MAG does not clearly identify vacant lands, and this should be modified. It appears that the following codes are the ones that suggest vacant land:

- Codes 0, 15: Unknown. This, of course, does not necessarily imply that these lands are vacant.
- Code 31: Vacant Desert.
- Codes 32 - 34: Agricultural. These lands are assumed to be eligible for development.

Future updates to the existing land use cover would improve the operation of the SAM if lands truly vacant were so designated. Infill development within the urban area is an especial concern.

Appropriate general plan designations for development were described in Chapters 5 (for general socioeconomic growth from DRAM/EMPAL) and Chapter 6 (for special population growth). Chapter 5 also described preliminary density levels that would be associated with each of the general plan codes. Readers will recall that residential densities were generally associated by cities with residential designations. Densities associated with employment related general plan uses however, were estimated based on countywide control totals. These densities were derived from Table 5-12 and Table 5-13.

Densities associated with downtown areas are especial concern. Downtown Phoenix, for example, clearly absorbs growth at much higher densities than do more suburban locations. Consequently, codes in the general plan should be modified to reflect this, so that higher densities can be associated with these uses.

8.3.2 Scoring Technique

A *hierarchical* scoring technique for designating the preference of developers for certain locations depending on the type of development was adopted.

The scoring technique is the same for all DRAM/EMPAL variables. For these, a three digit score *uhi*, where for each grid cell, was created for each digit reflects a different criterion:

- The hundreds digit *u* reflects *urbanization*, the theory being that cities grow outward and *"leap-frog"* development is discouraged by the lack of infrastructure or by public policy. The score is actually generated by measuring the distance between the cell and the closest developed property and is then normalized to a single digit score 0..9. Because this score is placed in the highest order (the hundreds) position, this criterion overshadows the other two.

- The tens digit *h* reflects *highway proximity*, the theory being that properties with visible highway access develop earlier than those without highway proximity. This score is created by generating 1-mile, 2-mile, and 5-mile buffers around all roads in the transportation network for the forecast year (assuming that all links in the network function as minor arterials or better.) This score is also normalized to a single digit score 0..9. Note that this method will force the land use forecasts to recognize the growth inducing impacts of new highways in the network.

- The units digit *I* reflects *infill*, the theory being that vacant properties next to developed land tend to develop earlier than those along vacant stretches of arterial streets. The number of cells immediately adjacent to the cell in question (there are 8 of them) which are developed, are counted, resulting in this score.

Note that this scoring technique considers three criteria; urbanization being the most important, highway proximity being second in importance, and infill being last in importance. The effect of this scoring technique should be that land close to the urban area will be developed first along major highways.

This same scoring technique for all DRAM/EMPAL variables: households as well as employment was used. New scoring techniques were required for special population groups, however, to reflect their unusual growth characteristics (discussed in Chapter 6):

- *Motels:* Motel growth occurs almost exclusively near freeway interchanges; therefore, the three digit score *xhu* is used, where *x* represents distance from the nearest freeway interchange, *h* represents highway proximity as before, and *u* represents the urbanization score as before.

- *Mobile Homes:* Mobile home/RV park growth occurs mostly in outlying Apache Junction. Therefore, a two digit score *ah* was created. The first digit, *a* represents whether the land is located in the general area around Apache Junction. The second digit represents highway proximity as before.

- *Nursing Homes:* Nursing homes are almost exclusively located northeast of Grand Avenue. Therefore, two criteria are considered to generate a two-digit number *pu*, where *p* represents proximity to existing nursing home and the urbanization digit is as before.

Other scoring techniques can be invented as well and can be integrated into the model when better notions of growth patterns are recognized.

8.3.3 Clustering

One would not expect growth of 1,000 households in a RAZ to be randomly dispersed through the desert — in fact households tend to grow as part of subdivisions. While we do not have good data on average development project size (or cluster size), we used 880 foot clusters to reflect all types of DRAM/EMPAL employment growth. This means that growth is allocated according to clusters 17.8 acres in size (for residential densities of 8 dwelling units per acre, this would imply subdivisions of 144 homes). Clustering does not mean land is always absorbed this way — other attributes connected with the land (score, eligibility) must be shared by each cell in the cluster.

Future research is required to provide better estimates of project size tenancies for different types of development. Random numbering to allow for random selection in case of tie scores.

8.3.4 Allocation

The allocation procedures merely involve sorting cells (and clusters) for each RAZ based on scores (in descending order). Growth associated with the RAZ is allocated to cells until it is fully absorbed. Two other conditions are:

- If there is not enough eligible land at the prescribed densities to absorb the growth forecast by DRAM/EMPAL, then densities will be increased across the board accordingly.

- If there is a cluster of land with scores that at build-out will exceed the RAZ control total, then it will be developed at a reduced density.

8.3.5 Features Not Implemented

Several SAM issues were not implemented because of the lack of any coverage or necessary information, but these coverages can be made and algorithms easily inserted into the framework of SAM:

- *Known Projects*

  Polygons associated with known projects for a given forecast year can be easily created and then gridded. The creation of a fourth digit to the scoring method, in the thousands place, was envisioned which would guarantee that these projects would rank first when it comes to allocation. In other words, the three digit *uhi* score discussed above would become a four digit *kuhi* score, where the *k* digit represents a *known* project.

- *Redevelopment*

  Redevelopment polygons can also be created and a mechanism inserted to remove existing development in existing cells underneath them.

- *Negative Growth*

  An operation to prorate reductions in housing or employment associated with a RAZ that experiences negative growth can be easily implemented.

8.4 SAM Implementation

The Subarea Allocation Model was implemented in ARC/INFO AML (and one custom-written c-program) and was installed in the directory ***/drive2/sam***. To run the application, the user types **&run sam** and responds to the menu interface.

8.4.1 Directory Structure

Figure 8-2 illustrates the directory structure associated with the sam subdirectory, whose contents are more fully described in Table 8-1.

8.4.2 Files

- *Input*

  The only input files to the process consist of the ***dfcst.dat*** and ***efcst.dat*** files created by the DRAM/EMPAL process. These are flat ASCII files with formats as shown in Table 8-2. These are placed in the subdirectory for the forecast.

- *Resource Files*

  The key resource files all reside in the ***/drive2/sam/data*** subdirectory and were described in Table 8-1. The two key files are a lookup table for existing land uses and a lookup table for general plan uses.

  Existing Land Uses: The ***/drive2/sam/data/exlup.info*** file contains a record for every existing land use code that is currently in use at MAG (and should be revised with changes in the existing land use cover). The field elig in this file flags those codes that are considered vacant. Users can alter the operation of the SAM model by redesignating new codes to be considered vacant.

Figure 8-2
Directory Structure

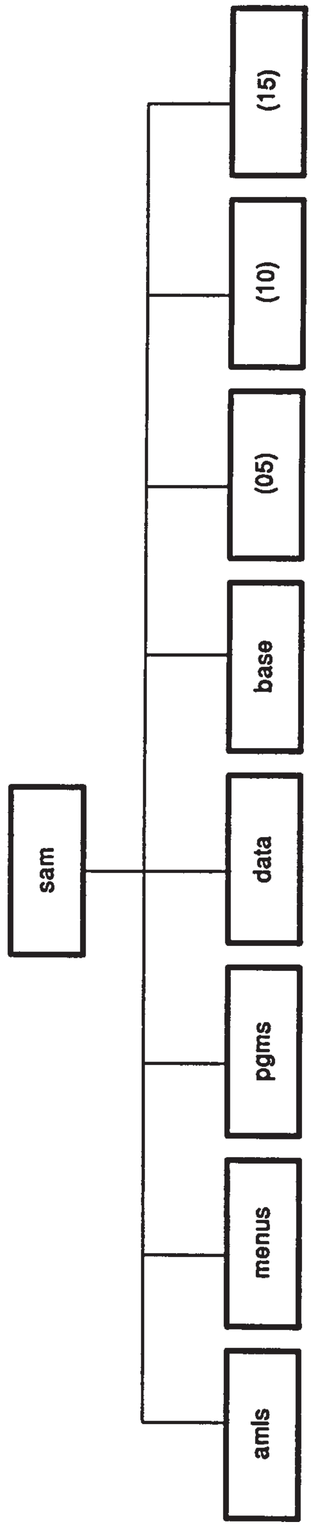


Table 8-1
Directory Structure Description

Subdirectory	Description
amls	Contains all of the aml routines which implement the SAM model.
menus	Contains all of the menus which appear in the SAM interface.
pgms	Contains source code for the magalloc.c program which is responsible for identifying grid cells to be developed. The actual compiled program resides in the parent directory and is known as /drive2/sam/magalloc.
base	Contains a number of datasets which describe base year land use (taken to be 1993). The base year land use cover is located here, along with its associated files. In addition, various point coverages for motels, nursing homes, etc. are located here (they are not used).
data	Contains a number of dataset resources which are used by the SAM application: • raz: The RAZ coverage • gplan: The general plan coverage • gplup.info: A lookup table describing the eligibility and density associated with each general plan code • exlup.info: A lookup table describing the eligibility of each existing land use code. • controlpop.info: A lookup table that gives countywide projections for each special population group, 1990 to the year 2040.
95, 05, etc.	Various subdirectories for each forecast year.

Table 8-2
Formats for the DFCST.ASC and the EFCST.ASC Input Files

File	Item	Width	Data Type	Description
dfcst.dat	RAZ	4	Integer	Official RAZ Number
	HHINC1	5	Integer	Projected Households
	HHINC2	5	Integer	by each income
	HHINC3	5	Integer	class
	HHINC4	5	Integer	
	HHINC5	5	Integer	
efcst.dat	RAZ	4	Integer	Official RAZ Number
	OTHEMP	5	Integer	Projected other employment
	GOVEMP	5	Integer	Projected government employment
	RETEMP	5	Integer	Projected retail employment
	OFFEMP	5	Integer	Projected office employment
	INDEMP	5	Integer	Projected industrial employment

General Plan Land Uses: The ***/drive2/sam/data/gplup.info*** file contains a record for every general plan code that is currently in use at MAG (and should be revised with changes in the general plan cover. This lookup table also flags what codes are eligible for what kinds of development.

Table 8-3 describes the format of these two important files.

- *Output*

  The results of the allocation process are represented by a series of grids, each representing a different land use sector, as shown in Table 8-4. These grids can then be used to generate a new growth polygon coverage and can be appended to the base year land use cover to create a detailed allocation data set for the forecast year.

8.4.3 Menus

The two main menus for the application are shown in Figure 8-3.

8.4.4 The magalloc.c Program

As has been described, the actual allocation process consists of rank-ordering cells in each RAZ with respect to their locational preference score. The allocation then descends this list, allocating growth according to whatever density is associated with the general plan code for the cell until all growth in the RAZ has been absorbed. This process is actually implemented by the custom-written c-program ***magalloc.c***. It therefore operates much faster than can be achieved by ARC/INFO AML.

The basic ARC/INFO data structure for grids consists of two files. The first is a matrix of values for each grid cell (actually compressed). These values point to a record in an INFO file called the ***.vat***. The ***.vat*** contains the real values associated with the cell. Therefore, an ASCII file can easily be dumped for a grid by dumping the ***.vat***; it can manipulate it, and can read the results back into the ***.vat***.

The ***magalloc.c*** program reads an ASCII version of the .vat which contains information on the RAZ number, the cluster number for the cell, the preliminary density associated with the cell, and its score. The ***magalloc.c*** program also reads an ASCII version of the growth table for the land use sector and computes the allocation. In fact, values in the ***.vat*** are tagged to indicate that they are the area to be developed. Also, new density values are computed, if necessary to totally absorb growth for the RAZ.

These results are output to a text file called ***alloc.dat*** and are then joined with the original ***.vat*** for the grid. The grid now reflects which cells have been "selected" for development.

Source code for the ***magalloc.c*** program has been installed in the directory ***/drive2/sam/pgms***.

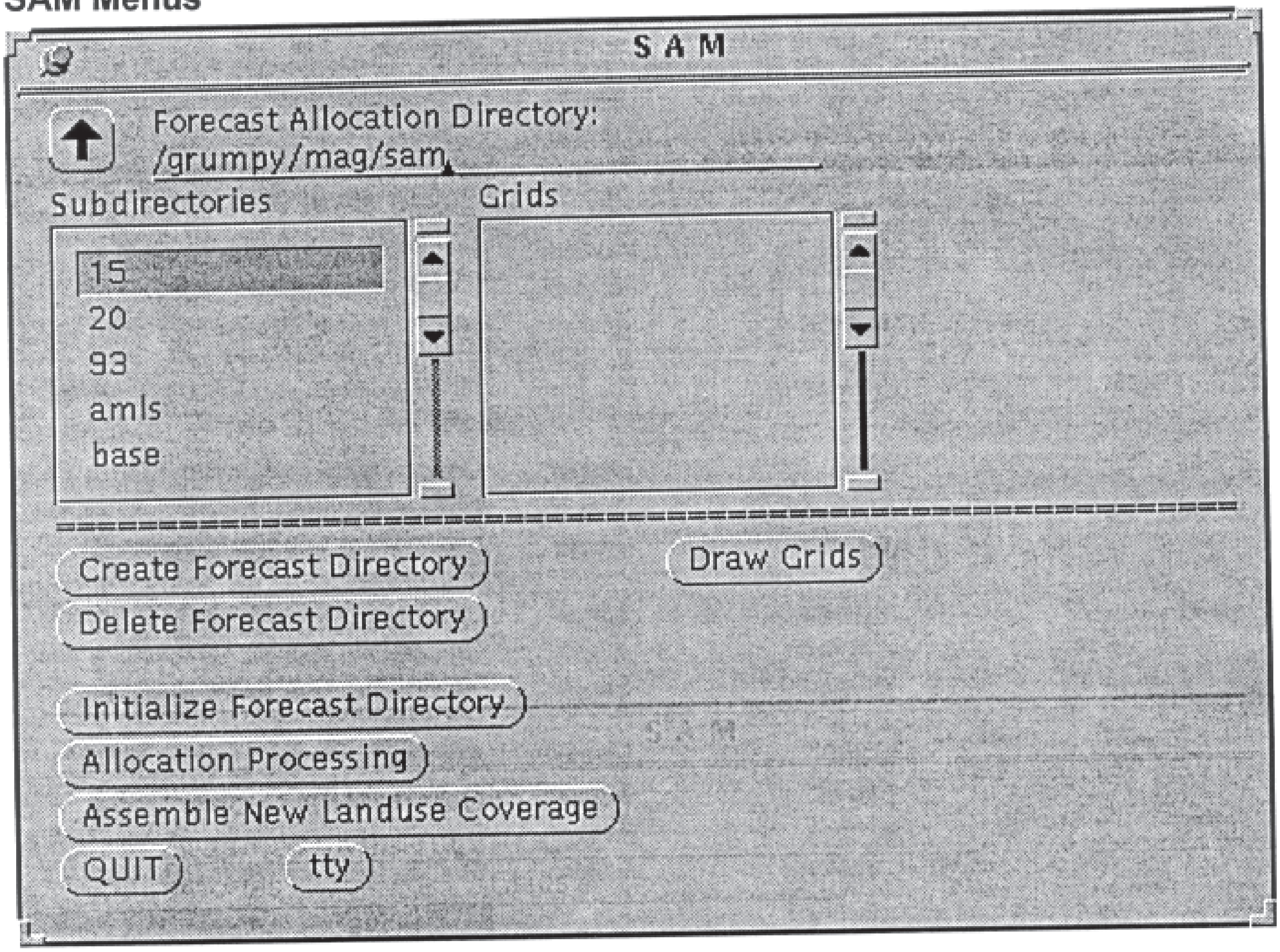
Table 8-3
The Two Lookup Tables

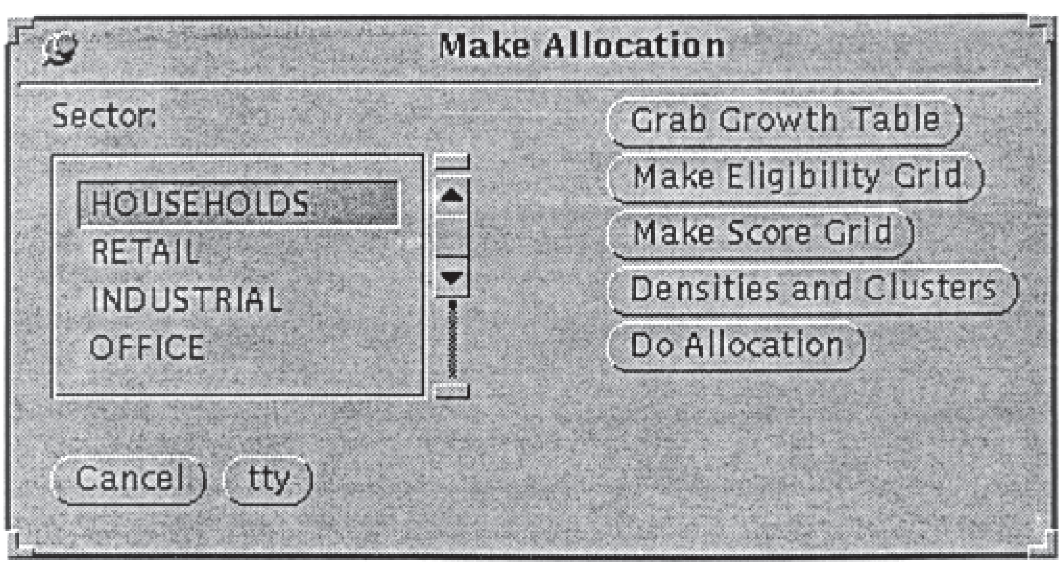
Lookup Table	Field	Data Type	Description
exlup.info	EXLU	integer	The existing land use code
	DESCRIPTION	character	Text description of the code
	ELIG	integer	Flags codes considered eligible for development (i.e., 1=vacant)
gplup.info	GPLU	integer	The general plan code
	DESCRIPTION	character	Text description of the code
	HHDENS	decimal	Residential density associated with the code
	EMPDENS	decimal	Employment density associated with the code
	HH	integer	Eligibility flag 1=ok for residential
	OTHEMP	integer	Eligibility flag 1=ok for other employment
	GOVEMP	integer	Eligibility flag 1=ok for government employment
	RETEMP	integer	Eligibility flag 1=ok for retail employment
	OFFEMP	integer	Eligibility flag 1=ok for office employment
	INDEMP	integer	Eligibility flag 1=ok for industrial employment
	NURSE	integer	Eligibility flag 1=ok for nursing homes
	DORM	integer	Eligibility flag 1=ok for dormitories
	MOTEL	integer	Eligibility flag 1=ok for motels and hotels
	MOBILEH	integer	Eligibility flag 1=ok for mobile homes/RV parks

Table 8-4
Growth Grids Created by SAM

Grid Name	Description
ggrowth_hh	Growth in households grid
ggrowth_oe	Growth in other employment grid
ggrowth_ge	Growth in government employment grid
ggrowth_re	Growth in retail employment grid
ggrowth_ofe	Growth in office employment grid
ggrowth_ie	Growth in industrial employment grid
ggrowth_nh	Growth in nursing homes grid
ggrowth_mh	Growth in mobile home/rv parks grid
ggrowth_m	Growth in motels/hotels grid

Figure 8-3
SAM Menus





8.5 Establishing the Base Year Coverage

Since the general approach to the SAM involves adding population and growth to a pre-existing land use coverage in terms of the land use developments associated with it, the SAM requires the establishment of a base year dataset upon which future forecasts will be added. Economic Strategies Group developed an initial version of an existing land use cover applicable to 1994, and this coverage provides the only basis for generating the SAM.

The existing land use cover ***exlu*** involves 13,224 land use polygons throughout Maricopa County, and provides considerably more detail about the distribution of land use and socioeconomic descriptors than does the TAZ structure, which involves 1,288 polygons. The only information about the polygons (other than area) is the land use code, which are described in Table 8-5. Note some of the idiosyncracies associated with the coding scheme — for example there are no codes specifically associated with retail.

Establishing a Base Year Dataset

An existing source of household and employment data needs to be identified in order to populate the 1994 existing land use coverage with socioeconomic information. A comparable dataset from the trip generation model would be best — that way it would be able to establish a baseline existing land use coverage dataset that is at least compatible with TAZ level datasets already in use at MAG. Several choices are available, including TAZ level trip generation datasets for 1990, 1993, and 1995. However, these issues were a concern:

- First, these datasets agree with current definitions of various employment types and were consistent with projections for them coming from DRAM/EMPAL.
- Second, these datasets do not carry estimates for households by income quintile (they contain only total households); therefore they must augmented with information created by DRAM/EMPAL.
- The 1993 dataset, which is closest to the baseline year of 1994, does not cover all of Maricopa County: it only covers the 1,272 TAZ modeling area. There is no comparable DRAM/EMPAL dataset for 1993, therefore income information will have to be taken from 1995.

Table 8-6 reports the compatibility between the various TAZ level trip generation datasets and those available from DRAM/EMPAL.

As noted in Table 8-6, thefears about basic incompatibilities between TAZ datasets used for trip generation and RAZ datasets from DRAM/EMPAL were unfounded. Therefore, the use of the TAZ data set for 1993 (***taz1993a.txt***) as a basis for populating the existing land use cover with attributes about households and employment. This file only represents the 1,272 TAZs which are internal to the modeling area. Therefore, projections of households and employment for the 16 external TAZs were developed by interpolating values in the DRAM/EMPAL projections for 1990 and 1995.

Table 8-5

Land Use Codes for the Existing Land Use Cover

1	Low Density Residential (SF)	23	Power Stations
2	Medium Density Residential (SF)	24	Railroads
3	High Density Residential (MF)	25	Airports
4	Mobile Homes/RV Parks		
5	Medium Residential Under Development	26	Freeways, Dams
		27	Parks
6	Low Density Commercial	28	Golf Courses
7	Medium Density Commercial	29	Lakes
		30	Rivers
8	Hotel, Resort	31	Vacant Desert
9	Regional Shopping Centers	32	Agriculture - Citrus
		33	Agriculture - Other Crops
10	Commercial Warehouse	34	Agriculture - Stockyards
11	Neighborhood Office Buildings	36	Canals
12	High-Rise Office Buildings		
		40	Nondevelopable - Other
13	Light Industrial	41	Nondevelopable - Forest
14	General Industrial	42	Nondevelopable - Mountain
		43	Nondevelopable - Gunnery Range
15	Unknown		
16	Institutions - Schools		
17	Institutions - Colleges		
18	Institutions - Universities		
19	Institutions - Small Hospitals		
20	Institutions - Large Hospitals		
21	Institutions - Public Facilities		
22	Institutions - Churches		

Table 8-6

TAZ/RAZ Datasets at MAG

	Variable	1990 TAZ	1990 RAZ	% Diff	1993 TAZ	1995 TAZ	1995 RAZ	% Diff
All	Other	117,972	117,972	100.0		129,986	129,108	99.3
1290	Government	130,194	130,195	100.0		170,393	170,394	100.0
	Retail	237,328	237,495	100.1		270,345	266,927	98.7
	Office	238,700	238,720	100.0		270,112	269,666	99.8
	Industrial	254,474	254,477	100.0		264,080	263,992	99.9
	Total	978,667	978,859	100.0		1,104,916	1,100,087	99.6
	Households	819,875	819,003	99.9		924,260	867,431	93.9
Model	Other	114,888	114,889	100.0	117,971	126,766	125,888	99.3
Area	Government	129,879	129,880	100.0	130,194	169,988	169,989	100.0
1272	Retail	235,183	235,350	100.1	237,328	267,776	264,358	98.7
	Office	237,824	237,844	100.0	238,700	269,076	268,630	99.8
	Industrial	252,202	252,205	100.0	254,474	261,733	261,645	99.9
	Total	969,976	970,168	100.0	978,667	1,095,339	1,090,510	99.6
	Households	813,396	812,561	99.9	875,585	917,048	860,655	93.9

Populating the Base Year Dataset

The intent is to populate polygons in the existing land use cover with TAZ-level information about household and employment; that is, to allocate office employment within a TAZ to office land use polygons in the TAZ. Right away, however, significant inconsistencies between the two databases were encountered. Numerous TAZs that contained office employment, for example, were found but contained no office land uses:

- There were 28 TAZs that contained office land use polygons, but the TAZ trip generation data set reports no office employment.
- There were 486 TAZs that contained office employment in the trip generation dataset, but they contained no office land uses.
- Therefore, 514 TAZs are inconsistent between the two datasets — about 40% of the region.

This inconsistency runs all the way through the two datasets for all land use and employment types. See Table 8-7.

Table 8-7
Consistency Between TAZ Projections and The Existing Land Use Cover

			TAZs with this Employment		
Employment	Type		None	Some	Total
Office	TAZs with this Land Use:	None	508	486	994
		Some	28	268	296
		Total	536	754	1290
Industrial	TAZs with this Land Use:	None	493	221	714
		Some	194	382	576
		Total	687	603	1290
Government	TAZs with this Land Use:	None	335	333	668
		Some	123	499	622
		Total	458	832	1290
Retail	TAZs with this Land Use:	None	162	260	422
		Some	28	840	868
		Total	190	1100	1290
Other	TAZs with this Land Use:	None	99	161	260
		Some	311	719	1030
		Total	410	880	1290

It would appear, then, that the information about land use patterns conveyed in the existing land use cover is different than that conveyed in the TAZ trip generation data sets. The goal was primarily to maintain consistency with TAZ level projections already in use at MAG, and therefore the 1993 TAZ level dataset was used as control totals for populating the existing land use cover.

The methodology could also take advantage of other data sources available from MAG. On the residential side, the parcel coverage yields counts of residential parcels in the region. On the employment side, the survey of major employers (50 employees or more) conducted by ESG in 1994 yields information on the location of employment. Therefore, the use of the data sources in association with TAZ level control totals and ignored conflicts with the existing land use cover.

Populating Existing Land Use Polygons with Households

The methodology for populating polygons in the existing land use cover with household counts was as follows:

- The parcel cover was overlaid on the existing land use cover, giving us a count of households in each polygon.
- The results were compared with TAZ level control totals:
    - Household counts were lowered across-the-board if they exceeded control totals for the TAZ.
    - Additional households were added to residential uses reflected in the existing land use cover if needed to match control totals for the TAZ. If there were no qualifying residential uses in the TAZ reflected in the existing land use cover, then households were added at an appropriate density for all developed uses.
- Age distributions for dwelling units were available from the parcel coverage. These were applied to existing land use polygons.

Populating Existing Land Use Polygons for Employment

The methodology for populating polygons in the existing land use cover with employment counts was as follows:

- The ESG survey of major employers was overlaid on the existing land use cover, giving us a count of employment in each polygon. Since the ESG survey did not classify these employment counts by land use type, the land use type was as specified by the land use polygon in question was assumed.
- Then, for each category of employment, the results were compared with the control totals for each TAZ.
    - Employment levels were reduced in cases where totals exceeded TAZ level controls.

- Employment was added to qualifying land use polygons at densities discussed in Chapter 5 if necessary to match TAZ level controls. If there were no polygons with the appropriate use, then any employment related polygon was used.

Populating Existing Land Use Polygons for Special Population Groups

The point coverages for nursing homes, mobil homes, motels, and institutions which were discussed earlier were all overlaid on the existing land use cover, so that polygons could be populated with data describing these uses.

Results

The results are a populated land use cover which has been installed in the directory ***/drive2/sam/base***. The database actually consists of the land use cover along with a set of related INFO files which address each component of land use.

The format of these files is described in Table 8-8.

8.5 Generating a TAZ Dataset

This section will describe the development of the application taz.aml which is responsible for generating a socioeconomic dataset for trip generation from any land use coverage and associated files. **What is important to note is that the taz.aml application generates a trip generation dataset for any TAZ coverage — it is not constrained to the existing MAG 1,288 zone TAZ coverage.**

General Approach

Table 8-9 describes the format of a typical socioeconomic dataset used by MAG trip generation models. Some of the variables which appear in this file come directly from the land use database created by the SAM, others are derived from basic variables or from geography,[5] as follows:

- Household data, as well as employment by each of the five employment classes, comes directly from the SAM land use cover. The target TAZ coverage is laid over the land use cover and statistics for these variables are aggregated.

- Population data can be derived from household data based on the population per household currently reflected for each TAZ in the existing TAZ datasets. That is, land use polygons inherit the population-per-household rate for the TAZ (and year) in which they reside.

---

[5] Also, some variables are presently unused (e.g., vehicle ownership)

Table 8-8
Existing Land Use Database Format

Datafile	Field	Format	Description
landuse<yr>	YEAR	I4	Forecast Year
	LU#	BI4	Land Use Polygon Identifier
	CODESCHEME	C1	E=existing land use coding scheme G=general plan land use coding scheme
	USE	I4	Land Use Code
	ACRES	F6.2	Total Acres associated with the polygon
demo<yr>	LU#	BI4	Land use polygon identifier
	HHINC1	I5	Resident households (Low Quintile)
	HHINC2	I5	Resident households (Low-Medium Quintile)
	HHINC3	I5	Resident households (Medium Quintile)
	HHINC4	I5	Resident households (Medium-High Quintile)
	HHINC5	I5	Resident households (High Quintile)
	HHAGE1	I5	Households (in dwelling units < 10 years old)
	HHAGE2	I5	Households (in dwelling units 10-19 years old)
	HHAGE3	I5	Households (in dwelling units 20-29 years old)
	HHAGE4	I5	Households (in dwelling units older than 29 years)
	TOTHH	I5	Total resident households (Total, all quintiles)
	TOTDU	I5	Total dwelling units
empl<yr>	LU#	BI4	Land use polygon identifier
	OTHEMP	I5	Other employment
	GOVEMP	I5	Government or Public employment
	RETEMP	I5	Retail employment
	OFFEMP	I5	Office employment
	INDEMP	I5	Industrial employment
	TOTEMP	I5	Total employment
mobileh<yr>	LU#	BI4	Land use polygon identifier
	TYPE	C1	M=mobile home; R=RV; B=both
	MHSPACES	I4	Mobile home units
	RVSPACES	I4	RV spaces
motel<yr>	LU#	BI4	Land use polygon identifier
	TYPE	C1	M = motel; H = hotel; R = resort
	ROOMS	I5	Number of rooms
nurse<yr>	LU#	BI4	Land use polygon identifier
	BEDS	I5	Number of beds
dorms<yr>	LU#	BI4	Land use polygon identifier
	TYPE	C1	J = jails; D = dorms; M = military
jails	BEDS	I5	Number of beds

Table 8-9
Trip Generation Dataset Structure

Field	Format	Description
YEAR	I4	Forecast Year
TAZID	I5	TAZ number
DISTRICT	I4	District number
MPA	2C	Metropolitan Planning Area Identifier
HHPOP	I6	Resident Population in Households
GQPOP	I5	Population in General Quarters
TRANPOP	I5	Transient Population
SEASPOP	I5	Seasonal Population
HH	I6	Resident Households
GQHH	I5	General Quarters Households
TRANHH	I5	Transient Population Households
SEASHH	I5	Seasonal Population Households
OTHEMP	I6	Other Employment
GOVEMP	I6	Government (or Public) Employment
RETEMP	I6	Retail Employment
OFFEMP	I6	Office Employment
INDEMP	I6	Industrial Employment
HHINC	I7	Median Household Income (1989 $)
DEVRESAREA *	I6	Developed Residential Area (square miles * 100)
UNDRESAREA *	I6	Undeveloped Residential Area (square miles * 100)
DEVEMPAREA *	I6	Developed Employment Area (square miles * 100)
UNDEVAREA *	I6	Undeveloped Area (square miles * 100)
AREA *	I6	Total TAZ Area (square miles * 100)
VEHICLES *	I6	Total Vehicles in TAZ
UPPERENROL	I6	Post High School Enrollment by Place of Residence
RETIRE	I2	Retirement Flag
ASU_DIST	I4	Distance to ASU (miles * 10)
DAILYPCOST	I4	Average Daily Parking Cost in TAZ (cents)
HOURLYPCOST	I4	Average Hourly Parking Cost in TAZ (cents)
ENPLANEMENTS	I6	Average daily enplanements at Sky Harbor
TERMTIME	I2	Terminal Time (minutes)

* Not used.

- Median income can be derived from the households by income class data obtained directly from the land use cover. This is done through a program to compare the TAZ household income distributions with those associated with various median incomes.

- Parking costs are obtained from the existing TAZ datasets. That is, parking costs associated with a target TAZ zone inherit the parking costs already coded in the 1,288 TAZ system. Terminal times are handled the same way.

- tranhh: Aggregated from the land use cover
- seashh: Aggregated from the land use cover
- othemp: Aggregated from the land use cover
- govemp: Aggregated from the land use cover
- retemp: Aggregated from the land use cover
- offemp: Aggregated from the land use cover
- indemp: Aggregated from the land use cover
- hhinc: Obtained by establishing the household income distribution for each target TAZ. With *all* call to the special c-program minc.c, the income distribution is compared with all of those in the region associated with every median income. A chi-square test establishes the closest fit.
- devresarea: Ignored: left blank
- undresarea: Ignored: left blank
- devemparea: Ignored: left blank
- undevarea: Ignored: left blank
- area: Computed from the area of the target TAZ polygon
- vehicles: Ignored: left blank
- upperenrol: Lookup table for college enrollments by site, assigned to the target TAZ polygon that contains the site.
- retire: Obtained by overlaying the target TAZ coverage with the original 1,288 zone TAZ coverage and setting the retirement flag accordingly
- asu_dist: Obtained by straightline airline distance measures from each target TAZ to the ASU coordinate.
- dailypcost: Obtained by overlaying the target TAZ coverage with the original 1,288 zone TAZ coverage and averaging daily parking costs associated with the original TAZs
- hourlypcost: Obtained by overlaying the target TAZ coverage with the original 1,288 zone TAZ coverage and averaging hourly parking costs associated with the original TAZs
- enplanements: Lookup table for enplanements for Sky Harbor for the applicable year, assigned to the target TAZ polygon that contains the airport.
- termtime: Obtained by overlaying the target TAZ coverage with the original 1,288 zone TAZ coverage and averaging terminal times associated with the original TAZs

More on Household Income

As indicated above, the method for estimating median household income from the input distribution of households by the five income classes is implemented by the minc.c program (source code for this is in the pgms subdirectory). The method is similar to that used in cross-classification trip generation models: a lookup table was devised which relates median income with an underlying distribution of households by income class; as you would expect, the distributions in TAZs with high median household incomes are skewed to the right. For any given input distribution, a chi-square test is performed for all possible median incomes

($0 to $150,000) to determine the closest fit. This establishes the median income for the zone.

MAG has RAZ level datasets which describe household income distributions (but not median income) and TAZ level datasets which describe median income (but not the underlying household distribution). Consequently, to construct the table, combined information from both the RAZ level DRAM/EMPAL dataset for 1993 and the TAZ trip generation dataset for 1993. Median incomes for RAZs were estimated from a weighted average of median incomes for the TAZs which are members, giving us 141 samples of RAZ household distributions and associated median income. A scatter plot relating median income with the percentage of low income households was generated and a curve was fitted (eyeballed). This was done also for each of the other income classes resulting in a lookup table that related median income with the percentage of households in each income class (the data was normalized to 100%).

Operation of the Application

The taz.aml application is easy to run: just type "&r taz" at the command prompt and a single onscreen form will appear requesting the names of the input files which are required (e.g., the land use cover name, the names of its associated files, and the TAZ cover for which a trip generation dataset is required). From there the application will run to completion.

8.6 Recommendations

This project resulted in a new framework for working with existing and projected socioeconomic data that offers MAG a number of benefits, including:

- The ability to work with any level of TAZ geography that is desired.
- The ability to automatically generate a trip generation dataset for that level of geography.

These capabilities substantially elevate the importance of the existing land use and general plan coverages: they not only serve MAG's purposes of displaying the distribution of land uses in the region, but with the emergence of the SAM, become *the* central database on which modeling is performed. In essence, the core geography for which socioeconomic and land use data is shifting from the TAZ system to these two land use coverages. This is consistent with the direction that MAG is taking in other data enhancement programs — which are becoming more comprehensive and more detailed in scope.

Accordingly, efforts to improve the accuracy and quantification of the land use cover and its linkage with DRAM/EMPAL and EMME/2 modeling are justified. Some of the recommendations toward that end include:

- The SAM model assumes that existing land use polygons designated as developed are, in fact, completely developed, and are therefore not available to accommodate growth without redevelopment projects. The extent to which the existing land use cover truly depicts this situation should be investigated. Possibly the development of an existing land use cover from a parcel coverage presents the best long term solution.

- Numerous conflicts between the existing land use cover and other sources of geographic data describing population and land use (e.g., locations of major employers depicted in the 1994 Major Employer Survey as well as mobile home parks, motels, nursing homes, etc.) were discovered. These differences should be resolved.

- The special census now underway, along with the acquisition of the Dunn-and-Bradstreet employer database, presents MAG with an excellent opportunity to repopulate the existing land use cover more accurately than could be done as part of this project.

- A number of different land use coding schemes have been used in the existing land use and general plan coverages. MAG should design and adopt a more consistent land use classification system that applies equally to these coverages.

- Cited a number of instances where employment definitions used by DRAM/EMPAL seem inconsistent with those used in the land use coverages (e.g., TAZs with office employees but no office land uses). This issue should be resolved, either by more closely aligning the definitions of employment associated with DRAM/EMPAL or by developing a methodology that allows say office employment in retail uses.

- Future planning efforts might also be devoted to developing a better understanding of business location decisions and adapting projection methodologies initiated in this project to that understanding.