Development of a Truck Model
For the San Diego Region

Final Report
December 2008

By:
PB Americas, Inc.

For:
San Diego Association of Governments (SANDAG)
Contents

1. Introduction ................................................................................................................... 3
2. A road map to freight modeling in San Diego County ................................................. 3
   2.1 Short-term options for freight modeling (0 to 5 years) .................................... 3
   2.2 Long-term strategies for modeling freight (5 to 10 years) ....................... 7
3. Data sources ................................................................................................................ 10
4. Model design ............................................................................................................... 11
5. Internal truck model .................................................................................................... 12
   5.1 Trip generation ............................................................................................... 12
   5.2 Special generators ........................................................................................... 13
   5.3 Trip distribution .............................................................................................. 14
6. External truck model ................................................................................................... 16
   6.1 FAF2 data ....................................................................................................... 16
   6.2 FAF2 adjustments ........................................................................................... 17
   6.3 Flow transformation and border crossings ..................................................... 18
   6.4 Assignment of external trips ........................................................................... 20
7. Temporal split and final assignment ........................................................................... 24
8. Model validation ........................................................................................................... 26
9. Truck model integration .............................................................................................. 31
10. Final remarks .............................................................................................................. 32
    10.1 Limitations ...................................................................................................... 33
    10.2 Recommendations for future truck counts...................................................... 35
List of figures

Figure 1: Scheme of the SANDAG Truck Model ............................................................ 11
Figure 2: SANDAG and SCAG regions ................................................................. 12
Figure 3: Trip distribution of regional flows ......................................................... 15
Figure 4: FAF2 data used by the external model ................................................. 17
Figure 5: FAF2 forecast and adjusted forecast .................................................... 18
Figure 6: Border flows to and from Mexico ........................................................ 19
Figure 7: FAF regions and counties................................................................. 20
Figure 8: County centroids for regional assignment ........................................ 21
Figure 9: Truck trips in 2002 assigned to the U.S. highway network .............. 22
Figure 10: Subarea analysis assignment ........................................................... 22
Figure 11: Hourly truck flows at WIM stations for an average weekday (AAWDT) 23
Figure 12: Entrance point labels ........................................................................ 24
Figure 13: Truck assignment for the AM Peak period ....................................... 25
Figure 14: Location of WIM stations ................................................................. 27
Figure 15: WIM count and model truck classifications ..................................... 27
Figure 16: Validation of simulated trucks at WIM stations ............................... 28
Figure 17: Validation of simulated trucks with Caltrans Counts .................... 30
Figure 18: Validation of simulated trucks with VOC Counts .......................... 30
Figure 19: Integration of the truck model with the SANDAG person travel model ... 32
Figure 20: Comparison of synthetic and actual assignment to identify important links 36
Figure 21: Significant links with volumes above the average ...................... 37
Figure 22: Significant links with volumes above the median ..................... 37
Figure 23: Subarea analysis settings ................................................................. 42

List of tables

Table 1: Truck classes of the Emissions Factors Model (Emfac2007) ................. 6
Table 2: Simulated truck types ............................................................................ 12
Table 3: Trip generation and attraction rates ..................................................... 13
Table 4: Friction factors for trip distribution by truck type .................................. 15
Table 5: Source of origins and destinations by flow direction .......................... 15
Table 6: Adjustment of FAF2 flows ................................................................. 18
Table 7: OD Matrix of subarea analysis for San Diego County ....................... 23
Table 8: Probabilities for time of day ............................................................... 24
Table 9: Simulated average trip length in miles .............................................. 26
Table 10: Simulated VMT by truck type ........................................................ 26
Table 11: Comparison of model fit with different count data .......................... 31
Table 12: File to specify adjustment of FAF2 forecast ...................................... 41
1. Introduction

The extent and impact of freight movements on urban freight networks are coming under increased scrutiny by a number of stakeholders. Transportation planners are concerned about their impact upon infrastructure and congestion. Civic leaders share the same concerns, although often equating efficient freight movement with economic competitiveness and market accessibility. The public has concerns about their safety, congestion, noise, and air quality impacts. Businesses and carriers wish to minimize the cost and uncertainty of freight transportation, and security officials highly desire increased visibility of freight content and network resiliency. All of these factors are germane in San Diego, which has a diverse economy and serves as an international land and sea gateway. The effort described in this report provides SANDAG with an urban freight model capable of addressing many of these issues.

While trucks share the same right of way as automobiles and exhibit some of the same operating characteristics, the underlying characteristics and drivers of freight demand are different and more complex than for person travel. An ideal freight model would be capable of handling many of these dynamics, which include the effect of multiple decision makers on trip making, the greatly increased use of trans-shipment facilities, heavy prevalence of multiple pickups and deliveries arranged into tours, and fleet and routing optimization. Unfortunately, such an ideal model does not yet exist in practice. The model described in this report, the first generation of freight models in San Diego, will put SANDAG on the path to the ideal model within the constraints imposed by budget, schedule, data availability, and the state of the art in freight modeling.

2. A road map to freight modeling in San Diego County

Like many regions across the country, freight has not been formally modeled in the San Diego region in the past. The options are somewhat limited in the short term, owing to the need to reach consensus on which key performance measures for freight are most important to policy-makers, the lack of experience with such models, the likely need to revisit the socioeconomic forecasting processes that support the modeling system, and the current lack of data on the behavior and patterns of the various agents that use and influence the freight transportation system. The following section describes the first phase of this project; review of existing freight modeling literature and a recommendation on short-term and long-term approaches for modeling freight in the San Diego region.

2.1 Short-term options for freight modeling (0 to 5 years)

Any approach to modeling freight in the short term will need to succeed with a minimum of local data on freight patterns or characteristics. SANDAG has few observed data on freight flows or behavior, and very limited vehicle classification count data on major
roadways and freeways within the region. This situation is common in most metropolitan areas across the country. The lack of locally observed data, however, does not preclude the development of data and models that will provide some insight, albeit incomplete, of freight movements within the region. Moreover, the implementation of such a model will help focus attention of the extent and location of data gaps about such movements, help illuminate freight issues within the larger transportation planning process, and guide the development of more sophisticated models.

There are two broad choices in modeling approaches open to SANDAG in the short term. Both are synthetic models that exploit successes achieved in other urban areas, and focus on commercial vehicle or truck movements in the absence of detailed freight data. The first is known as synthetic matrix estimation (SME), which attempts to identify the origin-destination matrix most likely to have given rise to observed truck counts across the urban area. It is useful to think of the SME process as the four-step sequential travel modeling process run in reverse, in that the outcome (flows on links) are known but the underlying patterns and behavior are only partially understood, if at all.

The process requires a seed matrix, which is a partially or fully observed trip matrix or the best guess at such. The technique iteratively converges to a solution that best matches the observed counts while deviating the least from the seed matrix. In theory these models should obtain a very high correlation between observed versus estimated flows, given that the model is heavily constrained by the former. In practice, however, the process sometimes fails to reach stability because of inconsistent or too few counts, having them in the wrong places, incorrect or incomplete information in the seed matrix, and the effects of network aggregation.

SME techniques have advanced steadily over the past two decades. The model described by Agnello, et al. (2004) is a good example of how such models are typically deployed in urban areas. It includes a process called adaptable assignment (Allen, 1997), which is somewhat of an abstraction of the SME process. It calculates scaled differences between observed and assigned volumes, and develops adjustment factors for affected origin-destination flows in the original matrix. The number of adjustments is based upon an iterative and interactive process, where the results from each iteration are evaluated by the analyst. This reduces the complexity of the process, and permits user insight into the working of what is often viewed as a “black box.” An important advantage of this technique is that it can be easily implemented using almost any transportation modeling package, obviating the need for specialized software.

Arguably the most advanced models of this type are those developed by List, et al. (2002) and List & Turnquist (1994), which overcomes many of the shortcomings noted above. They use a linear programming method that is more stable and efficient than the traditional maximum likelihood estimation used in most SME models, allow for the use of multiple pieces of data (in addition to the seed matrix) in estimation, and simultaneous solution for multiple vehicle classes. The latest version of this model has been successfully deployed in New York City (List, et al., 2002). Further work is underway to extend this model, which may make it attractive for further consideration.
The other option open to SANDAG in the short term is an adaptation of the four-step sequential modeling paradigm used for person travel analyses. The structure of these models and the typical values of their parameters are transferred from models implemented in similar urban areas, or are based upon national averages. The *Quick Response Freight Manual* (Cambridge Systematics, 2007) is used for the latter, and is based upon data and modeling practices from 21 urban areas. The QRFM distinguishes three truck types, namely 4-tire, 6+-tire single unit and 6+-tire combination unit commercial vehicles. Some of the data are relatively recent, while some date back twenty or more years. Information about the stability of these data and models over time is not available. A more recent survey of truck trip generation rates (Fischer & Han, 2001) provides updated information on trip generation characteristics than can be used in place of some of the older rates in the QRFM.

The QRFM includes three of the four steps found in sequential travel models. These include trip generation, trip distribution, and traffic assignment. Like their person travel model counterparts, these models also include exogenous estimates of external trips, and sometimes include special generators. The third step in sequential modeling, mode choice, is not typically carried out for a number of reasons. Trucking is the only viable mode of transport in almost all instances. In effect, the choice of mode is already known. Moreover, most of the important determinants in freight mode choice are not included in the modeling process (McGinnis, 1989). Determinants that decide over the choice of mode include goods value, goods perishability, goods weight, the costs of different modes and the distance. There are very few models that are able to reliably simulate mode choice for freight. In the short term, mode choice should be simplified by only considering the most important mode, which is truck travel.

Models based on the QRFM methodology are relatively easy and inexpensive to implement, although they have an uneven record of performance at best. Their correspondence with observed flow patterns is not as good as most person travel models unless augmented with a SME or similar process. However, this is due to the much higher variance in underlying behavior and lack of comprehensive data as much as theoretical or practical shortcomings. Thus, even a more elegant modeling approach would not likely provide more accurate results than a sketch planning tool like the QRFM until local data become available.

The recent update of the SCAG heavy duty truck model (Cambridge Systematics, 2008) is an important and useful extension to the knowledge base about truck travel in Southern California. Based partially on a trip diary survey of 362 trucks in the Los Angeles area, the model provides local data that are likely to be more representative of flows in the San Diego region than composite national data. This study derives factors for truck trip production and attraction based on employment by seven industries and number of households. In addition, the above-mentioned trip-generation overview by Fischer and Han (2001) as well as the *Quick Response Freight Manual* (QRFM) add further information about truck generation rates. Truck types in San Diego most likely shall be defined by the Emission Factor Model (Emfac2007) developed by the California Air...
Resources Board (2007: 7). Table 1 lists available sources for trip generation of every truck type of Emfac2007.

**Table 1: Truck classes of the Emissions Factors Model (Emfac2007)**

<table>
<thead>
<tr>
<th>Truck Type</th>
<th>Description</th>
<th>Weight Class (lbs.)</th>
<th>Sources for trip generation</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDT1</td>
<td>Light-Duty Trucks</td>
<td>0-3,750</td>
<td>QRFM, Fischer/Han</td>
</tr>
<tr>
<td>LDT2</td>
<td>Light-Duty Trucks</td>
<td>3,751-5,750</td>
<td>QRFM, Fischer/Han</td>
</tr>
<tr>
<td>MDV</td>
<td>Medium-Duty Trucks</td>
<td>5,751-8,500</td>
<td>QRFM, Fischer/Han</td>
</tr>
<tr>
<td>LHDT1</td>
<td>Light-Heavy-Duty</td>
<td>8,501-10,000</td>
<td>SCAG, QRFM, Fischer/Han</td>
</tr>
<tr>
<td>LHDT2</td>
<td>Light-Heavy-Duty</td>
<td>10,001-14,000</td>
<td>SCAG, QRFM, Fischer/Han</td>
</tr>
<tr>
<td>MHDT</td>
<td>Medium-Heavy-Duty</td>
<td>14,001-33,000</td>
<td>SCAG, QRFM, Fischer/Han</td>
</tr>
<tr>
<td>HHDT</td>
<td>Heavy-Heavy-Duty</td>
<td>33,001-60,000</td>
<td>SCAG, QRFM, Fischer/Han</td>
</tr>
</tbody>
</table>

As a starting point it is recommended that SANDAG implement a model using the QRFM methodology. The SCAG model provides a framework that has proved to represent truck traffic realistically in the Los Angeles region. Given the geographical proximity, estimated parameters are almost "local" for the SANDAG region. Depending on goodness-of-fit and availability of truck data in the future the model could be extended to a hybrid incorporating both QRFM methodology and an SME approach. This would be superior to either one individually, as the two approaches are more complementary than duplicative. The SME functionality in TransCAD could be coupled with a sequential travel model of the region based on SCAG and QRFM data. Adjustments to the QRFM parameters can be made to tailor the model to San Diego County, to the extent permitted by available data. This model complements the person travel model, using the same networks, zone structure, and zonal data. The person and freight models are joined in the traffic assignment step. There are several options for accomplishing this, with the optimal solution being a multi-class assignment that routes the flows from both models simultaneously.

This process allows information about changes in the socioeconomic variables used in the person model, economic trends and forecasts by sector, growth rates for trade and domestic freight from national models, likely industrial and commercial land development trends, and technological trends to be included in the forecasting process. The product of this step will be a set of growth rates by industry category and external sources that can add significant information to the forecasting process. In other words, building a freight model on the QFRM approach provides a flexible modeling tool that allows simulating alternative growth rates for population, employment and trade rates. This approach requires the collection of vehicle classification counts at several locations within the region, and on most functional classes of roadways. Indeed, such data are required irrespective of the modeling approach, for if a SME process was not employed the same data would be required for model validation. This hybrid model could achieve some degree of customization to the San Diego County, and provides the ability to test a limited range of alternatives for corridor and project evaluation. However, the model would be relatively insensitive to some policy issues, including price and cost sensitivities and changes in shipper and carrier behavior to response to changes in land use or transportation policies.
The Federal Highway Administration of the U.S. Department of Transportation has published the Freight Analysis Framework2 (FAF2), which provides goods flows in dollars and tons by commodity and mode (FHWA, 2002). Flows are given between 140 FAF2 regions, including 130 regions for the U.S. and 10 regions covering the rest of the world. Using average payloads these goods flows may be transformed into equivalent truck trips. Luckily, San Diego County is coded as a single FAF2 region (region code #9). The flows provided by FAF2 may serve to estimate truck trips that have either their origin or destination outside of San Diego County.

SANDAG provides data allowing to build a freight model, including socio-demographic data and truck counts by weight class. Further data about border crossing with Mexico addresses the unique location of San Diego between Mexico and the US West coast. More specific data needs should be established on the way of building and improving a freight model within the first five years.

From a current standpoint, there are three areas of data that appear particularly valuable to survey within the next few years. First, a special generator survey would help to identify and quantify truck generation and attraction by facilities that attract a particularly large number of trucks, such as warehouses or marine ports. And second, since SANDAG has highly detailed land use data readily available, it would be desirable to survey truck generation and attraction based on land use types. While employment at the TAZ level is available for 10 sectors only, SANDAG has spatial data for 94 land use categories at the MGRA level. A survey on truck generation/attraction by land use may increase the quality of model results significantly. Finally, data on truck trip generation at the various military installations within the county would provide useful data on both civilian and government vehicles leaving the bases. The limited data available for Air Force bases suggests that typical rates and diurnal patterns do not describe military facilities adequately.

2.2 Long-term strategies for modeling freight (5 to 10 years)

The establishment of a freight data program and expert panels will provide SANDAG with the better understanding of freight behavior and dynamics necessary to build and deploy more robust models. This in turn will help refine the data program and provide a broader view of the freight system to policy-makers. The better understanding of freight analysis requirements and data sources and limitations will facilitate the drafting of a long term modeling strategy for the region. Unfortunately, the state of practice has not evolved beyond that outlined in the interim model, although the literature abounds with extensions to and permutations of the approaches recommended above. However, none have yet proven more useful in practice than the recommended approach.

There are several important initiatives underway to improve the state of the practice in freight modeling. Many of these efforts are expected to produce better data and modeling tools for freight analyses within the next few years, which will coincide with the
development of the freight databases and maturing of the interim model. A second generation freight model for the region can be expected to have the following properties:

− **Explicit linkages to economic forecasts:** Freight flows are economic flows between producers and consumers of goods and services. The underlying motivation for freight is shaped by the structure and inter-relationships within the economy, which change more rapidly over time than the infrastructure that carries it. Linking freight flows to economic forecasts will provide a sound behavioral basis for freight demand and make the model more responsive to changes in economic conditions.

− **Translates commodity flows into vehicular flows:** Regional and national models express freight movements by commodity, which must be translated into loads on vehicles serving the region. An ideal model would be capable of tracking both commodity and vehicular flows (especially trucks moving within San Diego County).

− **Places San Diego in its global context:** Many of the flows within and through San Diego County have connections to other metropolitan areas in the country and to global trading partners. Changes in flows between San Diego and these partners can be appreciated by placing it within the larger geographical scale at which these flows occur.

− **Capture important dynamics:** There are several aspects of freight transportation that differ in important ways from person travel, to include multiple decision makers, much greater incidence of trip chaining and tour optimization, and the use of transshipment facilities. The latter have rapidly grown in importance over the past decade, and dominate in several markets serving urban areas, such as food and wholesale distribution.

− **Multimodal:** Freight moves by multiple modes, even within a metropolitan region. Flows into a regional distribution center may arrive by different mode than movements to final points of consumption, for example. The need to handle multiple modes is particularly important for San Diego County to capture links between sea freight, rail and trucks.

− **Robust truck-rail diversion analysis capability:** There is considerable interest at all levels of government in shifting freight from highway to rail modes. The pressures for such opportunities will increase as congestion increases in major urban areas and intercity corridors.

− **Policy sensitive:** The model must be responsive to the range of policy and investment options relevant to local decision-makers. Many technically elegant models languish unused because of their irrelevance to current issues facing transportation agencies.

− **Sparse data requirements:** Freight behavior and characteristics are much more highly diverse than person travel patterns. Higher sampling rates than for person travel surveys are required to accurately portray the flows. It is highly unlikely that traditional data collection approaches alone can fulfill the heavy data requirements imposed by traditional disaggregate travel models. Innovative techniques such as passive collection of ITS data must be employed to fill the gaps, which imply different model structures.

− **Operational:** The model must be capable of being successfully deployed in San Diego County based upon successful trials elsewhere. The software required to implement the model must be already available from previous work, reducing the risk and delays associated with software development and debugging.
Until the long-term model is clearly defined, it is difficult to identify data that are needed in addition to those used to develop the interim model. More likely, a long-term strategy would improve over the interim model by representing truck tours rather than trips (as mentioned in the fourth bullet point). If this is the case, a crucial piece of data required to develop such an improved model would be a survey of truck tours, gathering information about number and spatial distribution of stops on a truck tour, as well as the amount of freight carried from one stop to another.

The interim model will ideally be able to evolve into such a model. This can be accomplished by thinking of the ultimate model as a bi-level formulation, in which the upper level is based upon an economic modeling framework with connections to key trading partners, both domestically and abroad. San Diego County would be represented in this level of the model as a series of sub-county areas, and the flows would be expressed in commodity terms. This upper level would provide the control totals for the more detailed operational model at the lower level, and would be appropriately responsive to changes in commodity mixes, market changes, and new distribution patterns and technologies.

The lower level model might look very much like the interim model, except that its principal inputs would come from the upper level rather than being exogenously specified. The commodity mapping would extend to the lower level of the model, enabling the analyst to summarize flows by commodity, vehicle type, and shipper and carrier characteristics. The resulting bi-level model would replicate the important dynamics of freight using the measures most appropriate at each level. Moreover, it would allow the incremental development of freight models in San Diego County, growing up and out from the interim model to a fully behavioral model of freight and its interactions with other areas.

Unfortunately, it is probably beyond the abilities of the SANDAG to develop and maintain the upper level model on its own. The ability to construct such a model may well rest upon actions taken outside of the SANDAG, particularly the Freight Analysis Framework 2, or FAF2, (FHWA, 2002) and other national freight modeling initiatives, as well as statewide modeling that might be undertaken by the California Department Of Transportation (Caltrans). This will require a partnership with Caltrans and federal transportation agencies, for the upper level commodity and economic flows would logically emerge from a national or regional model. The Freight Analysis Framework 2 is presently only a prototype of such a system, requiring further development before it embodies the functionality and geographic detail required of the upper level model envisioned above. Fortunately, the Federal Highway Administration is seeking to regain a leadership position in freight modeling, and has plans to extend the FAF2 in important ways. It is expected that the redesigned FAF2 and other tools will likely be widely available in the next several years, presenting the San Diego County with many options for collaborative modeling that are not available today.
3. Data sources

Developing a truck model for San Diego County requires a reasonable amount of input data. As an aggregate transportation model, this SANDAG truck model limits data requirements significantly. Nevertheless, there are a number of datasets that are required as a minimum to build a reasonable truck model.

One of the main input data to the truck model is employment and number of households by TAZ. Both were provided by SANDAG, and distinguish ten employment categories (Agriculture/Mining, Construction, Government, Manufacturing, Military, Other Services, Retail, Self-Employment, Transportation/Warehousing/Utilities, Wholesale), which are aggregated by the model to seven categories (Agriculture/Mining/Construction, Retail, Government, Manufacturing, Transportation/Utilities, Wholesale, Other). A distance skim matrix is used in truck trip distribution and reflects travel routes due to off-peak levels of congestion based on both car and truck traffic in the off-peak period.

Special generators are sources of truck flows that can not be easily identified by their employment. Cruise ship terminals, for instance, create much more truck traffic than their employment would suggest. Data for special generators include number of cruise ship berthed in the past five years as published by the Port of San Diego. For the special generator airport, the Caltrans has published total amount of cargo shipped through Californian airports in the last ten years.

For the external truck model, the Freight Analysis Framework 2 (FAF2) as published by the Federal Highway Administration (2002) is used. The FAF2 data are provided in four different data sets that are all explored for the external model:

- Domestic: Commodity flows between domestic origins and destinations in short tons.
- Border: Commodity flows by land from Canada and Mexico via ports of entry on the U.S. border to domestic destinations and from the U.S. via ports of exit on the U.S. border to Canada and Mexico in short tons.
- Sea: Commodity flows by water from overseas origins via ports of entry to domestic destinations and from domestic origins via ports of exit to overseas destinations in short tons.
- Air: Commodity flows by air from abroad origins via airports of entry to domestic destinations and from domestic origins via airports of exit to abroad destinations in short tons.

To validate the model only limited data were available. The most reliable data sources are counts at Weight in Motion (WIM) stations. These provide hourly counts over many days throughout the year. 15 vehicle types classified by axles are distinguished. Ten WIM stations were available for this model development. No weight measurements were available. Caltrans provides counts of all major highways across the State. Unfortunately, these locations are recounted only every six years, making the counts less reliable for
validation. Finally, Vehicle Occupancy and Classification Survey Data (VOC) were available for 106 locations all over San Diego County, distinguishing trucks by axles.

4. Model design

The vast majority of personal trips are of short distance of 50 miles or less. In contrast, truck trips may be a few miles only or may be several thousand miles long crossing the entire country. While a fairly detailed level of simulation can be achieved for San Diego County, it is unfeasible to handle truck trips at the same level of detail for the entire US. At the same time, only within the local study area a high level of detail is of interest. Whether a truck trip from San Diego goes to Berkeley or Oakland is irrelevant for analyses within San Diego County. A multi-layer model approach allows simulating internal trips and long-distance trips with a different richness in detail appropriate to each.

To account for two different levels of geographic scale, the truck model was developed as two separate modules (Figure 1). An internal truck model generates local truck trips and an external truck model simulates truck trips across the entire US. The internal model creates truck trips that stay within San Diego County (internal-internal or II). The external model generates flows between all counties in the US. Thus, for San Diego County it simulates trips that leave the county (internal-external or IE), trips that enter the county (external-internal or EI) and flows that go through the county (external-external or EE). The latter are mostly truck trips that go into or originate in Mexico.
The internal model is a trip-based model with trip generation based on socio-economic data, trip distribution and a temporal split. A special generator module simulates trips from sources that cannot be captured by their employment only, such as cruise ship terminals or airports. The external model uses FAF2 data, disaggregates flows to counties, temporarily assigns these flows to the external network, and provides external trips by three time periods. Flows of both models are combined in a multi-class traffic assignment that assigns flows from the internal and external truck models in conjunction with flows from the person transport model. Finally, these flows will used to estimate air quality.

5. Internal truck model

Internal truck trips within the San Diego County are simulated by the internal model. It is based on a model developed by Cambridge Systematics, Inc. (2008) for the Southern California Association of Governments (SCAG), called SCAG model in this report. Given the geographic proximity (Figure 2) it is assumed that truck generation patterns are comparable in San Diego County. Trucks are distinguished by three truck types defined by their weight class (Table 2). These truck types are compatible with vehicle types defined by Emfac2007, which is used by SANDAG to calculate vehicle emissions. The internal truck model generates a matrix of internal-internal flows that later is combined with external truck trips for final assignment.

<table>
<thead>
<tr>
<th>Type</th>
<th>Weight class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Light-Heavy Duty</td>
<td>LHD</td>
</tr>
<tr>
<td>Medium-Heavy Duty</td>
<td>MHD</td>
</tr>
<tr>
<td>Heavy-Heavy Duty</td>
<td>HHD</td>
</tr>
</tbody>
</table>

The internal truck model is developed in TransCAD, Version 4.8. A GISDK script was written to easily rerun the model.

5.1 Trip generation

Employment and household data are used to estimate the number of truck trips that originate in or are destined to a given zone. Employment is given by ten categories at the TAZ level. Households are given at the MGRA level (with 33,353 polygons) and are aggregated to TAZ (4593 zones plus 12 external stations).
The SCAG model provides truck generation and attraction factors for the three truck types by 7 region types (such as CBD, suburban, rural). As a comparable classification type does not exist in San Diego County, the average of these regions was used. An input file <\data\regionType.csv> is provided in case it should be decided in the future to introduce a different region type for every TAZ. Table 3 summarizes the trip production and attraction rates for each truck type. The last row shows which socio-demographic variables of the SANDAG database have been used to calculate the actual number of trucks originating and ending in every zone.

Table 3: Trip generation and attraction rates

<table>
<thead>
<tr>
<th>Type</th>
<th>Households</th>
<th>Ag/Min/Constr</th>
<th>Retail</th>
<th>Government</th>
<th>Manufacturing</th>
<th>Transp/Utilities</th>
<th>Wholesale</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>LHD Production</td>
<td>0.01950</td>
<td>0.06644</td>
<td>0.15371</td>
<td>0.00400</td>
<td>0.03530</td>
<td>0.17889</td>
<td>0.03930</td>
<td>0.00910</td>
</tr>
<tr>
<td>LHD Attraction</td>
<td>0.01950</td>
<td>0.07010</td>
<td>0.15541</td>
<td>0.00400</td>
<td>0.03530</td>
<td>0.14629</td>
<td>0.03930</td>
<td>0.00910</td>
</tr>
<tr>
<td>MHD Production</td>
<td>0.01100</td>
<td>0.02827</td>
<td>0.04043</td>
<td>0.03181</td>
<td>0.05750</td>
<td>0.05210</td>
<td>0.06500</td>
<td>0.01410</td>
</tr>
<tr>
<td>MHD Attraction</td>
<td>0.01100</td>
<td>0.03267</td>
<td>0.04133</td>
<td>0.03270</td>
<td>0.05750</td>
<td>0.09454</td>
<td>0.06500</td>
<td>0.01410</td>
</tr>
<tr>
<td>HHD Production</td>
<td>0.00166</td>
<td>0.03894</td>
<td>0.01694</td>
<td>0.00719</td>
<td>0.03910</td>
<td>0.05530</td>
<td>0.06330</td>
<td>0.00300</td>
</tr>
<tr>
<td>HHD Attraction</td>
<td>0.00324</td>
<td>0.04124</td>
<td>0.02214</td>
<td>0.00433</td>
<td>0.03910</td>
<td>0.04617</td>
<td>0.06330</td>
<td>0.00300</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SANDAG variables applied</th>
<th>HH</th>
<th>Agriculture And Mining, Construction</th>
<th>Retail</th>
<th>Government</th>
<th>Manufacturing</th>
<th>Transportation</th>
<th>Warehousing Utilities</th>
<th>Wholesale</th>
<th>Other Services, Self-Employment</th>
</tr>
</thead>
</table>

5.2 Special generators

The module *Special gen* (short for special generator, see Figure 1) creates trips that cannot simply be derived from either employment or land use, but amount to a significant number of truck trips. Special generators specified in this model include military sites, the cruise ship terminal and the airport. An example for the exogenous input file with special generators is shown in Appendix A.

Military sites produce a significant amount of truck trips, both for food supply, machinery supply and training purposes and military purposes. No specific truck generation factor for militaries is given in the SCAG model, those would be treated as "Other Employment". Because their specific truck-generation rate and the large number of 85,070 employees (in 2004), military personnel are treated as special generators. As no truck count data for military sites was available, an average of 1 daily truck per 50 employees was assumed.

The cruise ship terminal is treated as a special generator because the relatively small number of employees working at the terminal hides the large number of trucks that serve the port when a cruise ship anchors. From 2003 to 2007, an average of 177 cruise ships per year stopped in San Diego. For an average weekday, this translates into 0.51 cruise ships per day. In lack of better data, an average of 20 trucks per cruise ship was assumed.

The third special generator considered in this model is the San Diego International Airport. There are three major truck-generating sources at the airport. First, there are passengers that shop and dine at the airport. These trucks are generated in the general
truck-generation process (section 5.1), as the airport with more than 5,000 employees generates over 1,000 trucks per day. Second, the freight terminal produces truck trips, as most if not all goods flown into San Diego are transported by truck to the final destination. These truck trips are handled by the external truck model (see section 6), which generates truck trips from and to all U.S. airport based on commodity flows. Thirdly, mail is flown into and out of San Diego airport. Since mail is not classified as commodity, the external model does not capture those trips. A special generator "Airport Mail" takes account of such truck trips. The California Department of Transportation (2008) provides tons shipped per year for every airport (San Diego in 2002: 167,344 tons). FAF2 reports 5 tons for 2002, thus it is assumed that there are 167,339 tons of mail going through San Diego airport per year. Assuming every mail truck carries 5 tons, there are 33,468 trucks per year. Divided by 365.25 days per year and multiplied with a factor of 1.048 to transform from an average day to an average week day, there are 96 trucks per day serving the airport with airmail. Since it is assumed that trucks carry goods both to and from the airport, there are 48 mail trucks in each direction per day.

These truck trips produced and attracted by special generators are added to the truck trip table provided by the general truck generation step (as described in section 5.1). In line with the three truck types used by the model, trucks produced by special generators need to be split into LHD, MHD and HHD trucks. The exogenous file with special generators allows specifying the share of each truck type. As default, a model-wide average is assumed.

After calculating total truck production and attraction (based on employment and special generators), the values need to be balanced. Creating a distribution matrix (as described in section 5.3) requires that the sum of origins equals the sum of destinations. A procedure that uses the average of the two sums is applied to balance productions and attractions. The output is a list of zones with the number of trucks that have their origin and number of trucks that have their destination in a given zone.

### 5.3 Trip distribution

Trips generated by the modules *Trip generation* and *Special gen* are fed into the *Trip distribution* module. The task of this step is to link origins with destinations by creating actual trips. The trip distribution module applies a gravity model, which distributes trips based on the number of origins and destinations for every zone as well as an impedance between zones:

\[ V_{i,j} = P_i \cdot A_j \cdot \exp\left(-\alpha \cdot t_{i,j}\right) \]

where

- \( V_{i,j} \) : Truck volume from \( i \) to \( j \)
- \( P_i \) : Production in \( i \)
- \( A_j \) : Attraction in \( j \)
- \( \alpha \) : Friction factor by truck type
- \( t_{i,j} \) : Off peak travel time from \( i \) to \( j \)
The friction factor $\alpha$ is defined for every truck type by the SCAG model (Cambridge Systematics 2008: 6-14 to 6-15). The applied values are shown in Table 4. Since internal-external and external-internal mostly are medium-heavy and heavy-heavy duty trucks, their friction factor is used for these long-distance trips. There is no friction factor for external-external trips since their origin and destination is defined by the external model, and no further distribution within San Diego County is necessary.

<table>
<thead>
<tr>
<th>Flow type</th>
<th>Origin</th>
<th>Destination</th>
<th>Balancing</th>
</tr>
</thead>
<tbody>
<tr>
<td>II Internal to Internal</td>
<td>Truck Production based on employment</td>
<td>Truck Attraction based on employment</td>
<td>Balanced to the average of the two sums</td>
</tr>
<tr>
<td>IE Internal to External</td>
<td>Truck Production based on employment</td>
<td>IE trips of external model</td>
<td>Balanced to sum of IE trips of external model</td>
</tr>
<tr>
<td>EI External to Internal</td>
<td>EI trips of external model</td>
<td>Truck Attraction based on Employment</td>
<td>Balanced to sum of EI trips of external model</td>
</tr>
<tr>
<td>EE External to External</td>
<td>EE trips based on external model</td>
<td>EE trips based on external model</td>
<td>No balancing required as trips are balanced</td>
</tr>
</tbody>
</table>

Table 4: Friction factors for trip distribution by truck type

<table>
<thead>
<tr>
<th>Truck type</th>
<th>Friction factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Light-heavy duty</td>
<td>0.045</td>
</tr>
<tr>
<td>Medium-heavy duty</td>
<td>0.03</td>
</tr>
<tr>
<td>Heavy-heavy duty</td>
<td>0.03</td>
</tr>
<tr>
<td>Internal-external</td>
<td>0.03</td>
</tr>
<tr>
<td>External-Internal</td>
<td>0.03</td>
</tr>
</tbody>
</table>

The simulation of external trips is described in section 6. The output provided by the external model requires a different procedure of trip distribution. Three different directions of travel are provided by the external model (Figure 3): External-external (EE or through trips), internal-external (IE) and external-internal (EI).

For EE trips, no trip distribution is required, since origin and destination are set. IE and EI trips, however, need to be distributed. For IE trips the destination (external exit station) is set by the external model, but the origin within San Diego County is unknown. Thus, a destination-constrained trip distribution is needed. Since no information is available where an IE trips starts, the number of truck productions per zone is used as the best proxy. The sum of truck production, obviously, is much larger than the sum of all IE trips. By normalizing the productions within San Diego County to the sum of all IE trips, the sum of the origins is made the same as the sum of the IE flows. For EI flows, the reverse is done, and truck attractions within San Diego County are normalized to the sum of EI Flows and used at destinations. Error! Reference source not found. gives an overview of origins and destinations by flow direction.

Table 5: Source of origins and destinations by flow direction
6. External truck model

The external model is based on freight flows given by the Freight Analysis Framework, Version 2 (FAF2) published by the Federal Highway Administration (2002). The FAF2 provides freight flows in tons or dollar value between 140 FAF regions all over the US for the year 2002 plus forecasts from 2010 to 2035 in five years increment. The mode is distinguished as well as 43 commodities. Based on commodity-specific payload factors (Battelle 2002: 29), these flows are transformed into annual truck trips. Based on assumptions about the empty-load rate and an average number of business days per year, daily truck trips are estimated between the 140 FAF regions. A national average of truck type distribution is used to distinguish three truck types defined for the internal truck model.

Truck trips between 140 FAF regions are further disaggregated to truck trips between 3,241 counties. These truck flows are assigned to a U.S. highway network. This assignment is only used to extract the number of trucks that enter or leave San Diego County at any given highway exit, including through trips and trips that have their origin or destination in San Diego County. As a result, the External Model generates internal-external (IE), external-internal (EI) and external-external (EE) trips.

The external model is written in the Java programming language, and is treated as an exogenous input for the internal model. It is assumed that policies affecting truck trips in San Diego County do not influence truck trips in the rest of North America. Thus, the external model can run one time, produce all necessary input files for the internal model, and there is no need to rerun the external model for different policy scenarios. However, the external model can be run with different assumptions about truck trip growth rates in North America. Appendix B provides a manual how the external model is run with changed assumptions about growth of truck trips.

6.1 FAF2 data

The FAF2 data contain different modes and mode combinations. For the purpose at hand, only the mode 'Truck' was used. Figure 4 shows data included and excluded in this analysis. Combinations such as 'Truck & Rail' or 'Air & Truck' were omitted assuming that the longer part of that trip is done by Rail or Air, respectively, and only a small portion is done by truck. As the data do not allow distinguishing which part of the trip has been made by which mode, combined modes were disregarded completely. 'Air & Truck (International)' was included as these allow extrapolating the portion done by truck from the international airport to the domestic destination, and vice versa. Of the 200,320 flows that are omitted, only a fairly small portion of these trips are by truck. The shortcoming is assumed to be acceptably small. Border data were considered with the portion from the border crossing to the domestic destination or from the domestic origin to the border crossing. Likewise, sea and air freight was included as a trip from or to the domestic port or airport.
The FAF2 forecast of commodities flows reaches until the year 2035. SANDAG is interested in simulations up to the year 2050, therefore, future years need to be estimated. A plot of total FAF2 flows between 2002 and 2035 revealed that the forecast assumes an exponential growth of commodity flows (thin blue line in Figure 5). This growth rate appears fairly optimistic. Extrapolating this growth beyond the year 2035 potentially overestimates truck trips significantly. One reason to question this steep growth is that the forecast assumes more or less constant gas prices beyond 2006 (Battelle 2007: 3), an assumption that appears to be unlikely given the limited supply with crude oil and a historically continuously growing demand. Furthermore, an adjustment of future years appears reasonable since economic outlooks that were available at the beginning of this century were more optimistic than current growth assumptions.

As the growth from 2002 to 2020 is relatively constant, these values are used without any adjustment. The growth rate from 2002 to 2020 is extrapolated into the future and used for the years 2025 to 2050 (red line in Figure 5). This more conservative growth rate is assumed to more likely resemble future growth in truck traffic. Table 6 summarizes which FAF2 flows serve as the basis for every forecast year and which have been adjusted.

6.2 FAF2 adjustments

The FAF2 forecast of commodities flows reaches until the year 2035. SANDAG is interested in simulations up to the year 2050, therefore, future years need to be estimated. A plot of total FAF2 flows between 2002 and 2035 revealed that the forecast assumes an exponential growth of commodity flows (thin blue line in Figure 5). This growth rate appears fairly optimistic. Extrapolating this growth beyond the year 2035 potentially overestimates truck trips significantly. One reason to question this steep growth is that the forecast assumes more or less constant gas prices beyond 2006 (Battelle 2007: 3), an assumption that appears to be unlikely given the limited supply with crude oil and a historically continuously growing demand. Furthermore, an adjustment of future years appears reasonable since economic outlooks that were available at the beginning of this century were more optimistic than current growth assumptions.

As the growth from 2002 to 2020 is relatively constant, these values are used without any adjustment. The growth rate from 2002 to 2020 is extrapolated into the future and used for the years 2025 to 2050 (red line in Figure 5). This more conservative growth rate is assumed to more likely resemble future growth in truck traffic. Table 6 summarizes which FAF2 flows serve as the basis for every forecast year and which have been adjusted.
### Appendix B provides a manual how the regional model is run to produce truck trips under different assumptions of truck trip growth.

#### 6.3 Flow transformation and border crossings

FAF2 flows between 140 FAF regions are given in short tons (= 2,000 lb or 907.18 kg). The FAF2 also provides payload factors that may be used to transform commodity flows in tons into truck trips. Some trucks will run empty and a vacancy rate of 19.4 percent (U.S. Census Bureau 2008) was added to all flows. In line with the internal truck model (section 5) the yearly FAF2 flows are converted to flows of an average weekday. Flows were divided by 365.25 days and a factor of 1.048 was estimated based on national truck

### Table 6: Adjustment of FAF2 flows

<table>
<thead>
<tr>
<th>Year</th>
<th>Flow Patterns</th>
<th>Adjustment</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>FAF2 2002</td>
<td>none</td>
</tr>
<tr>
<td>2010</td>
<td>FAF2 2010</td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>FAF2 2015</td>
<td></td>
</tr>
<tr>
<td>2020</td>
<td>FAF2 2020</td>
<td></td>
</tr>
<tr>
<td>2025</td>
<td>FAF2 2025</td>
<td>normalized to match growth rate 2002 to 2020</td>
</tr>
<tr>
<td>2030</td>
<td>FAF2 2030</td>
<td></td>
</tr>
<tr>
<td>2035</td>
<td>FAF2 2035</td>
<td></td>
</tr>
<tr>
<td>2040</td>
<td>FAF2 2040</td>
<td></td>
</tr>
<tr>
<td>2045</td>
<td>FAF2 2045</td>
<td></td>
</tr>
<tr>
<td>2050</td>
<td>FAF2 2050</td>
<td></td>
</tr>
</tbody>
</table>

*no FAF2 forecast beyond 2035 exists*
data to transform Annual Average Daily Traffic (AADT) into Annual Average Weekday Traffic (AAWDT).

**Border Crossings**

It is common to under-estimate truck flows across the border. Cross-border traffic follows different rules than general truck traffic. Many truck trips drive a short distance only. Particularly at the Mexican-US Border, numerous maquiladoras produce many truck trips with a very short distance. Maquiladoras are factories that import materials tax-free, process these materials and re-export the final product. According to the website of the City of San Diego (2008), there are more than 570 maquiladoras in Tijuana, which provide about 100,000 jobs. These maquiladoras produce a large number of cross-border truck trips, most of them reaching only a short distance into Mexico as well as into the U.S. Experts report empty-truck rates of up to 50 percent for cross-border traffic, which is significantly higher than the national average of about 20 percent (U.S. Census Bureau 2008). Furthermore, the mix of truck types may be different for trucks crossing the border than for domestic truck flows. Because of its geographical location it is crucial to produce the right number of border-crossing trucks.

A first assignment without any adjustments produced truck trips at the two border crossings Otay Mesa and Tecate that were about 45 percent below the counts provided by the USDOT. The two border crossings in neighboring Imperial County, Mexicali and Andrade, had a very similar rate of under-estimating truck trips. A manual adjustment was required to account for a higher share of empty trucks and a different fleet composition. A factor of 1.898 was calculated that is used to increase all border-crossing truck flows on Californian border stations. Subsequently, border crossings at all four Californian border stations match well with truck counts.

The FAF2 data does not only define the amount of goods shipped across the border, but also names the FAF region in which the flows crossed the border. For flows from Mexico to the U.S. that cross the border in San Diego County, there are two possible border stations: Otay Mesa and Tecate (the crossing San Ysidro at the southern end of I-5 is for

![Figure 6: Border flows to and from Mexico](image)
cars only). The U.S. highway network was extended by a few links into Mexico as shown in Figure 6. Flows from Mexico begin where shown by the red arrow. The assignment process defines whether trips enter through Otay Mesa or Tecate. This way, trucks may shift from Otay Mesa to Tecate if congestions at Otay Mesa produces excessive delays. Though no delay at the border is simulated (it is implicitly assumed that all border crossings have the same delay time), congestion on roads in the U.S. may trigger some route change. The same system has been set up for the two entrances (Mexicali and Andrade) into Imperial County.

**Disaggregation to Counties**

FAF regions resemble major business areas. In less urbanized areas such as Maine or Montana, entire states comprise a single FAF region. California is subdivided into five FAF regions: San Diego County, Los Angeles Metropolitan Area, San Jose/San Francisco/Oakland Metropolitan Area, Sacramento/Arden-Arcade/Truckee Area, and the remainder of the state (Figure 7). To increase spatial resolution, the flows between FAF regions are disaggregated into flows between 3,241 counties in the US based on total employment per county. This is based on the assumption that most truck trips are generated and attracted by businesses rather than households. For San Diego County it is of particular interest to distinguish traffic that originates in Los Angeles and Orange County from trips that originate in the remainder of the Los Angeles FAF region. Given the large area of the Los Angeles FAF region, a finer spatial resolution of flows is desirable to assign truck trips to the right entry and exit points of San Diego County.

### 6.4 Assignment of external trips

In a first attempt, the geographical centroid of every county was used for assignment (left map in Figure 8). In the big picture of the 48 lower States this appears to be an acceptable simplification. Looking into more detail, however, the geographical centroid rarely matches the activity center within a county.

To improve the overall picture, the largest city (in terms of population) within every county was selected and the county centroid was moved to the centroid of the largest city in every county (right map in Figure 8). In San Diego County, for instance, the county centroid was moved from a remote location in the Cleveland National Forest to the geographical centroid of the City of San Diego.
Even though this does change flows within every county quite a bit, the adjustment is expected to have only minor impact on the number of trucks that cross from one county into another. As the external model is only used to provide the number of IE, EI and EE trips at all highway entrances of San Diego County, the impact of shifting the centroid is small. For visualization purposes, however, it is desirable to locate the centroid in a reasonable position.

For the assignment, a daily capacity of every highway link had to be estimated. In lack of true data the capacity was estimated based on the highway class and the number of lanes. While Interstate Highways (both Urban Interstate and Rural Interstate) are assumed to have a capacity of 2,400 vehicles per hour per lane (vphpl), all other highways are assumed to have a capacity of 1,700 vphpl. The daily capacity is assumed to be ten times higher than the hourly capacity, as most transportation demand arises during daylight hours.

Since trucks are not the only vehicles on the streets, autos need to be added as background volume on the highway network to create the right level of congestion. In rural areas, a Level of Service (LOS) C is assumed, such that auto volumes corresponding to a volume-to-capacity (V/C) ratio of 0.6 filled by cars. This is assigned to highways classified as "Rural Interstate", "Rural Major Collector", "Rural Minor Arterial", "Rural Minor Collector", "Rural Principal Arterial" or "Unknown". In urban areas, highways are assumed to be more congested, and highways are expected to operate between LOS D and E, using a V/C ratio of 0.9 that is filled by cars.

Figure 9 shows truck flows assigned to the U.S. highway network of the lower 48 States. The assignment for the year 2002 shows the expected pattern with a high concentration of truck flows between New York, Chicago and Houston as well as San Francisco and Los Angeles. The data needed for the region covered by the internal truck model requires a subarea analysis of U.S. truck flows. Figure 10 shows a map San Diego County in light yellow and external truck flows in blue.
Table 7 shows data that is used from the subarea analysis in the internal truck model. From every highway entrance into the study area the matrix shows how many daily trips stay within the area and how many trips go to all the other entrance points. These values are passed on to the truck distribution of the internal truck model (section 5.3). Figure 12 locates all entrance points shown in Table 7 with a red dot. Note that San Diego County is represented by one point only (yellow dot in Figure 12). This is needed for the nationwide assignment of the external model. Flows that go to "San Diego County" are split into different destination TAZ within the county by the internal truck model, as described in section 5.3. Smaller roads that cross the outline of San Diego County did not
receive any truck traffic in the assignment. These smaller roads are insignificant for external truck traffic. At the border crossing "Otay Mesa" only one entrance point (Otay Mesa East) is open for truck traffic.

Table 7: OD Matrix of subarea analysis for San Diego County

<table>
<thead>
<tr>
<th>From</th>
<th>To</th>
<th>I-5</th>
<th>I-15</th>
<th>I-8</th>
<th>Marine Port</th>
<th>Airport</th>
<th>Otay Mesa</th>
<th>Tecate</th>
<th>San Diego</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>I-5</td>
<td></td>
<td>348.43</td>
<td>5.84</td>
<td>567.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2652.57</td>
</tr>
<tr>
<td>I-15</td>
<td></td>
<td></td>
<td>3.13</td>
<td>103.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1127.25</td>
</tr>
<tr>
<td>I-8</td>
<td></td>
<td>230.03</td>
<td>2.36</td>
<td>120.58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>408.80</td>
</tr>
<tr>
<td>Marine Port</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>68.71</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>264.73</td>
</tr>
<tr>
<td>Airport</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>350.97</td>
</tr>
<tr>
<td>Otay Mesa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.74</td>
</tr>
<tr>
<td>Tecate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>77.90</td>
</tr>
<tr>
<td>San Diego</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1661.93</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>2153.88</td>
<td>735.35</td>
<td>621.10</td>
<td>21.66</td>
<td>0.68</td>
<td>1137.35</td>
<td>120.58</td>
<td></td>
<td>5310.45</td>
</tr>
</tbody>
</table>

Figure 11: Hourly truck flows at WIM stations for an average weekday (AAWDT)
The internal truck model defines the three truck types LHD, MHD and HHD to be compatible with the Emfac2007 model, which estimates emissions. The external truck model is unable to distinguish truck types. Before adding external truck trips to the distributions matrices of the internal truck model, external trucks need to be split into three truck types. In lack of more reliable data, a nationwide average of 30.7 percent for LHD, 15.5 percent for MHD and 53.8 percent for HHD have been assumed.

7. Temporal split and final assignment

The person travel model developed by SANDAG distinguishes three time-of-day periods: morning peak (6:00 to 8:59), afternoon peak (15:00 to 17:59) and off-peak (remaining hours). To assign trucks with cars jointly, the distribution matrices (as described in section 5.3) are split into these three time-of-day periods.

Obviously, no true information about the time of day distribution for every OD is available. The data of the WIM stations, however, do report truck flows by the hour. Figure 11 shows flows at the available WIM stations for every hour, proposing an average of 17 percent in the morning peak, 15.6 percent in the afternoon peak and the remaining hours carry 67.4 percent.

However, looking more closely there are three curves that show a different shape, namely WIM stations 607, 666 and 687. These are WIM stations located on the border. As the borders are closed for truck traffic at night, the curves are almost flat from 8 pm to 5 am. Therefore, it was decided to use different time-of-day distributions that come from or go to a border crossing.

Table 8 gives the share for each time-of-day period. Any flow that has its origin or destination at a border crossing gets a higher share in the morning and afternoon peak and a smaller share in the off-peak period, as the border is closed at night.

The OD matrices for truck trips for the three periods of time are assigned to the SANDAG highway network in a multi-class assignment. Besides the existing five
personal vehicle classes drive-alone no toll (dan), drive-alone toll (dat), shared ride no toll no HOV (s2+nn), shared ride no toll HOV (s2+nh) and shared ride toll HOV (s2+th), the three truck classes light-heavy duty (LHD), medium-heavy duty (MHD) and heavy-heavy duty (HHD) are distinguished, making a total of eight cores that are assigned simultaneously. In line with the person travel model, a logit delay function is used for the assignment.

Due to the vehicle length and a larger distance between every vehicle, trucks use more space on the highway. Therefore, passenger car equivalents (PCE) are used to account for this increased space requirement. A PCE of 1.3 for LHD, 1.5 for MHD and 2.5 for HHD is used in the assignment. Furthermore, trucks are excluded from any HOV link. Figure 13 shows the final assignment of all truck trips (LHD + MHD + HHD) in the am peak period.

A true calibration was not required. It was decided that the application of SCAG model parameters was the best option available at this point of time. Any adjustment of these parameters to local conditions in San Diego County was avoided due to lack of specific data. The most promising option for adjusting parameters would be survey data, which was not available for San Diego County. Fortunately, the validation shown in the following section indicates, that an adjustment of parameters is not immediately required.
Table 9 and Table 10 summarize simulated trips length and simulated vehicle miles of travel (VMT). Three truck types are distinguished, as well as Internal and External. Internal covers all trips that have their origin and destination within San Diego County. External covers trips that have at least one end outside San Diego County, but the numbers represent only that portion of the trip that happened within San Diego County. The actual distance traveled by these external trips is significantly larger. VMT is further subdivided into the three time-of-day periods. This distinction has not been given for trip length, as the average distance is identical in all three time periods.

### Table 9: Simulated average trip length in miles

<table>
<thead>
<tr>
<th></th>
<th>Internal</th>
<th>External</th>
</tr>
</thead>
<tbody>
<tr>
<td>LHD</td>
<td>17.84</td>
<td>48.37</td>
</tr>
<tr>
<td>MHD</td>
<td>20.25</td>
<td>48.37</td>
</tr>
<tr>
<td>HHD</td>
<td>20.37</td>
<td>48.37</td>
</tr>
</tbody>
</table>

### Table 10: Simulated VMT by truck type

<table>
<thead>
<tr>
<th>AM peak period</th>
<th>PM peak period</th>
<th>Off peak period</th>
<th>All periods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal</td>
<td>External*</td>
<td>Internal</td>
<td>External*</td>
</tr>
<tr>
<td>LHD</td>
<td>208,009</td>
<td>207,951</td>
<td>187,591</td>
</tr>
<tr>
<td>MHD</td>
<td>171,215</td>
<td>171,169</td>
<td>154,408</td>
</tr>
<tr>
<td>HHD</td>
<td>77,426</td>
<td>77,414</td>
<td>69,898</td>
</tr>
<tr>
<td>Total</td>
<td>456,650</td>
<td>456,534</td>
<td>411,897</td>
</tr>
</tbody>
</table>

* includes IE, EI and EE trips only with the portion of the trip within San Diego County

## 8. Model validation

The validation step serves to test if the model results are close to observed truck volumes, which would indicate that the simulated truck volumes are considered to be a valid representation of real-world traffic flows. The reliability of traffic counts the model is compared to is crucial. If a link has been counted only one time, the reliability of this count is considered to be poor. Too large is the possibility that the counted truck volume does not represent an average day, flows could be distracted by a construction site, an accident or a special event on that particular day. It is not uncommon that single day counts deviate significantly from an average day. In San Diego County, standard deviations of up to 57 percent of the average value have been found. Therefore, it is essential to carefully check the validity of counts when comparing model results to traffic counts.

The most reliable data found is Weigh-In-Motion (WIM) data. These truck volumes are counted daily and given by hour of the day and truck class (defined by axles), counted on many days throughout the course of the year. Figure 14 shows the eight WIM locations with San Diego County. They cover a significant part of the Interstate highway system within the county. A weight distribution, which would be valuable to validate different truck types, unfortunately was not available.
The available classification of WIM counts by axles makes it difficult to compare these counts with simulated truck volumes by weight class directly. Figure 15 shows the total number of trucks at the eight WIM stations in San Diego County. The first column contains trucks as classified by the WIM data. Some classes are grouped together, as those groups of classes are considered to be the closest equivalents to the simulated truck classes. The second column shows the sum of simulated volumes at the eight WIM sites, classified by weight class as defined in the truck model. There is little resemblance between the groups of the WIM data and the simulated truck classes. WIM classes 9-14 are much larger than class HHD, classes 7-8 have a much smaller volume than class MHD, and classes 5-6 are larger than class LHD. The reason for this discrepancy is that one data set is classified by axles, and the other one by weight. WIM class 9 with 5 axles, for instance, will have some trucks with more than 33,000 lbs and some trucks with less. For this reason it is impossible to precisely relate the WIM truck classification with simulated truck classes.

The definition of trucks, however, has been set to be compatible with the Emfac2007 model, which is used to calculate emissions. The Emfac2007 model requires truck by weight class, as weight is a much better indicator for emissions than number of axles. The
simulated truck classes are fully compatible with the Emfac2007 classifications. The size of the Emfac2007 column in Figure 15 is for visualization purposes only and does not indicate that the Emfac2007 model expected exactly these volumes.

One might think that the total number of simulated truck (SANDAG column in Figure 15) should match the total number of truck counted at the WIM stations. Unfortunately, another inconsistency prohibits this comparison. While the model includes pick-up trucks and vans with a weight of 8,500 lbs or higher as light-heavy duty (LHD) trucks, the WIM counts do not specify heavy pick-up trucks and vans separately. The WIM class 3 contains all pick-up trucks and vans, regardless of their weight. The total number of class 3 vehicles at the eight WIM stations is 210,000 vehicles. Only the small part of pick-up trucks and vans with a weight of 8,500 lbs or higher are included in the simulation of LHD trucks.

Figure 16 compares the simulated volumes with counts at WIM stations in each direction. The majority of volumes are slightly overestimated, which is in line with the observation explained in Figure 15 that the simulation includes pick-up trucks and vans over 8,500 lbs, whereas the sum of WIM classes 5-14 excludes such vehicles. Interstate 805 is noteworthy underestimated, which is assumed to be due to some local truck generation not explained by employment numbers. Due to the limit comparability of truck types LHD, MHD and HHD with WIM classes (compare Figure 15), it is unreasonable to validate truck types individually.

Validating a model against eight stations is limited, as the fit on other links potentially
could be quite different. For this reason it is desirable to check the model results against further data. The Caltrans publishes counts on major highway links every year. There are 159 stations in San Diego County, and commonly one sixth of those is recounted every six. Since a couple of count locations have been abandoned over time, only 43 of these counts have been updated between 2000 and 2006. Out of these 43, 17 were actually recounted, while 26 have been estimated based on surrounding counts and assumptions about traffic growth.

Figure 17 shows the relatively poor fit between simulated truck volumes and the Caltrans counts. These results, however, have to be seen with caution. Only 17 of these points have been counted within the last six years, which by itself is a fairly long period of time. Equally as in the WIM data, Caltrans defines trucks by number of axles, while the model simulates vehicles of 8,500 lbs or more. Furthermore, this is a one-point-in-time count. Special circumstances, such as accidents, construction sites, sport events, to name just a few, could have altered the volumes. An analysis of the WIM counts revealed a standard deviation of 15 percent of the average annual weekday count. Single counts may deviate substantially from an annual average weekday volume.

In search for more truck count data, SANDAG was able to provide additional Vehicle Occupancy and Classification Survey Data (VOC). This survey was designed as a one-day count to estimate the number of vehicles by type (classified by axles) and the occupancy of passenger vehicles. 106 sites (19 freeway sites and 87 local road sites) all over San Diego County were counted within one week in September 2000 starting at 6:30 am and ending at 2:00 pm.
Figure 18 compares simulated volumes with truck counts at VOC sites. The fit is relatively poor, which may be due to a couple of reasons. First, the counts have been made for half a day, allowing only to compare the morning am peak period. The model defines this period as 6:00 am to 9:00 am, the counts reach one hour less from 6:30 am to 8:30 am. For this reason it is expected that the model results are higher, as they cover three instead of two hours. Again, these count data are classified by axles and not by weight class. The definition of trucks in the VOC counts seems to be wider than in the model, as on one HOV link were counted 63 trucks, while the model does not allow any trucks on HOV lanes. And finally, the VOC counts again are a one-point-of-time count, capturing all special circumstances that may have been present on the count day, while the model aims at simulating an annual average weekday. For these reasons, the comparison with VOC counts allows limited conclusions about the model fit.

Figure 18 compares simulated volumes with truck counts at VOC sites. The fit is relatively poor, which may be due to a couple of reasons. First, the counts have been made for half a day, allowing only to compare the morning am peak period. The model defines this period as 6:00 am to 9:00 am, the counts reach one hour less from 6:30 am to 8:30 am. For this reason it is expected that the model results are higher, as they cover three instead of two hours. Again, these count data are classified by axles and not by weight class. The definition of trucks in the VOC counts seems to be wider than in the model, as on one HOV link were counted 63 trucks, while the model does not allow any trucks on HOV lanes. And finally, the VOC counts again are a one-point-of-time count, capturing all special circumstances that may have been present on the count day, while the model aims at simulating an annual average weekday. For these reasons, the comparison with VOC counts allows limited conclusions about the model fit.

Figure 18: Validation of simulated trucks with Caltrans Counts

Figure 18 compares simulated volumes with truck counts at VOC sites. The fit is relatively poor, which may be due to a couple of reasons. First, the counts have been made for half a day, allowing only to compare the morning am peak period. The model defines this period as 6:00 am to 9:00 am, the counts reach one hour less from 6:30 am to 8:30 am. For this reason it is expected that the model results are higher, as they cover three instead of two hours. Again, these count data are classified by axles and not by weight class. The definition of trucks in the VOC counts seems to be wider than in the model, as on one HOV link were counted 63 trucks, while the model does not allow any trucks on HOV lanes. And finally, the VOC counts again are a one-point-of-time count, capturing all special circumstances that may have been present on the count day, while the model aims at simulating an annual average weekday. For these reasons, the comparison with VOC counts allows limited conclusions about the model fit.

compares the three data sets WIM, Caltrans and VOC. Several measurements were tested to describe how well count data and model results match. As has been explained above, all count data sets do not fit very well to validate the model. The available data has been explored as far as possible. The best fit can be found between the WIM data and the model results. However, due to the small number of WIM count stations, the model fit has to be taken with care until further validation becomes possible. More count data would be desirable to truly validate the SANDAG truck model.
Table 11: Comparison of model fit with different count data

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Equation</th>
<th>Desirable value</th>
<th>WIM counts</th>
<th>Caltrans counts</th>
<th>VOC counts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of counts</td>
<td>$\sum^N$</td>
<td>high</td>
<td>17</td>
<td>156</td>
<td>104</td>
</tr>
<tr>
<td>Correlation R</td>
<td>$\sqrt{\frac{N \cdot \sum (P \cdot O) - \sum P - \sum O}{\sum P^2 - (\sum P)^2 \sum O^2 - (\sum O)^2}}$</td>
<td>1</td>
<td>0.886</td>
<td>0.816</td>
<td>0.929</td>
</tr>
<tr>
<td>Correlation R²</td>
<td>$R^2$</td>
<td>1</td>
<td>0.785</td>
<td>0.666</td>
<td>0.863</td>
</tr>
<tr>
<td>Root Mean Square Error (RMSE)*</td>
<td>$\sqrt{\frac{\sum (P-O)^2}{N-1}}$</td>
<td>small</td>
<td>1,421.7</td>
<td>3,215.2</td>
<td>179.5</td>
</tr>
<tr>
<td>Percent Root Mean Square Error (%RMSE)</td>
<td>$\frac{RMSE \cdot 100}{\bar{O}}$</td>
<td>0</td>
<td>34.29</td>
<td>59.71</td>
<td>242.33</td>
</tr>
<tr>
<td>Index of agreement (IA)</td>
<td>$1 - \frac{\sum (P-O)^2}{\sum (P - \bar{P}) + \sum (O - \bar{O})}$</td>
<td>1</td>
<td>0.904</td>
<td>0.883</td>
<td>0.838</td>
</tr>
<tr>
<td>Modeling efficiency (ME)</td>
<td>$1 - \frac{\sum (O - \bar{O})^2 - \sum (P-O)^2}{\sum (O - \bar{O})^2}$</td>
<td>0</td>
<td>0.447</td>
<td>0.575</td>
<td>1.452</td>
</tr>
</tbody>
</table>

$O$: Observed, $\bar{O}$: Mean of observed, $P$: Predicted, $\bar{P}$: Mean of predicted, $N$: Number of counts

(* use of 1 fewer degrees of freedom to account for a small number of observations)

It would be desirable to validate the average trip lengths reported in Table 9. Unfortunately, such data are not available for San Diego County. The average trip lengths are in line with those surveyed in the SCAG region. Simulated average truck trips lengths in San Diego are approximately 60 percent smaller than surveyed trip lengths in the SCAG region, which is in line with the difference in area covered by the two regions (compare Figure 2). Admittedly, this is a qualitative assessment only.

9. Truck model integration

Before this truck model was developed, the person travel model generated truck trips in a simplified way. The Other-Other (OO) trip purpose included trips made by heavy trucks. To integrate the truck model with the person travel model, OO trip rates had to be reduced in order to prevent the double-counting of heavy trucks. The reduction in OO trip ends accounts for the longer trip length of truck trips compared to OO trips, such that the total VMT caused by truck trips is correctly removed from the OO purpose.

The OO trip rates were reduced by first computing a trip rate factor based on the ratio of internal-internal truck trip length (20 miles) to OO trip length (6 miles), or 3.25. In other words, each truck trip is worth 3.25 OO trips in VMT reduction. Next, the OO trip productions and attractions generated by land-use category were compared to truck trip productions and attractions by land-use category. A target reduction in OO productions and attractions was calculated by land-use category, using the factor of 3.25 to account for trip length differences, so that the truck trip ends would be properly accounted for.
spatially. The ‘rates’ input file was adjusted accordingly and the model was re-run to verify that the calculations were implemented correctly. In total, Other-Other trips were reduced by approximately 385,000 trips. Most of the reductions occurred in truck-intensive land-use categories such as heavy industry.

The final integration of the truck model with the person model is fairly straightforward. The SANDAG person travel model works in four steps: Model Inputs, First Stage, Feedback Loop and Final Stage (Figure 19). Highway skims generated in the First Stage are fed into the truck model to run truck generation and truck distribution. The truck trip tables are then fed into the Highway Assignment in the First Stage. The third step Feedback Loop improves skim matrices based on actual travel demand. In the fourth step Final Stage, updated skims are used to rerun truck distribution and updated truck trip tables are fed into the final multi-class highway assignment.

10. Final remarks

In urban areas, about 6 percent of all vehicles are trucks. On some highway links in more rural areas, the share of trucks may be as high as 25 percent. As an example, the model simulates 5 percent trucks on S-78, while the share of truck on I-8 close to the border with Imperial county reaches 20.2 percent in the AM peak period. Including truck trips into a person travel model completes the picture by simulating all vehicle classes endogenously. The truck model that has been developed for SANDAG allows testing a wide range of policy options, such as restricting trucks on certain routes or raising toll for trucks at a different level than for cars.
Overall, simulation results are expected to improve. Previously, truck trips were estimated by simply increasing trip rates for non home-based (NHB) trips. All NHB trips needed to go through a mode choice step, and as a result a few of these feigned truck trips were using public transport. This is not uncommon as a first step, however, it is desirable to overcome this simplification as SANDAG did with this truck model development. Having a proper truck model allows simulating a more realistic reaction of road users on policies that are tested in scenarios. A certain scenario may affect person travel strongly and truck trips negligibly. Such a distinction may only be modeled by simulating personal vehicles and trucks separately.

It is worth noting that the entire truck model has been developed without k-factors. Many models use k-factors to tweak the model output to match counts in the base year. Though this greatly improves the model fit in the base year, k-factors have a doubtable effect on future traffic flows. As the k-factor commonly is static and not influenced by changes in the simulation system, future simulation periods tend to be synthetically bound to the base year traffic patterns. This also means that models with a heavy k-factor use tend to be less responsive to policy scenarios. As part of the model output is based on static k-factors, the representation of behavioral changes in the model is limited. For this reason it was valued high to avoid k-factors in the development of this truck model.

It turned out to be a fruitful approach to build a multi-layer model. Having a coarse model for external truck trips and a fairly detailed model for internal truck trips allows simulating every task at an appropriate level of detail. Whereas it is of great interest for SANDAG whether a truck trip ends in Chula Vista or Imperial Beach, it is not important for the purpose at hand to distinguish a truck going to northern Seattle or southern Seattle. Working at different levels of detail allows being efficient on the one hand and providing detail where desired on the other hand.

### 10.1 Limitations

By definition, a model is a simplified representation of reality. Accordingly, a model is not able to replicate reality in every aspect. Addressing simplifications and possible shortcomings of the model frankly allows the user understanding model output and assessing the appropriateness of the model design for certain tasks.

The model is fairly limited in simulating the effect of changed wait times to cross the border with Mexico. Due to limited information about truck origins and destinations in Mexico, the truck model is unable to model shift of flows between different border crossings reliably.

There are two levels of inaccuracy at border crossings. First, whether a truck from Mexico crosses the border in San Diego County or Imperial County is fixed by the FAF2 forecast. The FAF2 data define how many goods cross the border in the two Californian counties on the Mexican border, both for the base year and for any future simulation period. In reality, some trucks currently crossing the border in Tecate in San Diego
County may shift to Mexicali in Imperial County if traffic conditions in San Diego get severely congested. The model, however, is not designed to simulate such a shift. Second, the distinction between the two border crossings within San Diego County, Otay Mesa and Tecate, is fairly crude. Due to the lack spatial representation in Mexico the travel time from the zone "Mexico" to either border crossing in San Diego County is identical. The only reason why a truck trip may choose one border crossing over the other is based on the travel time to its final destination. As a result, all trucks coming from Mexico through FAF2 region San Diego and going to San Diego County or any county north thereof are simulated to use Otay Mesa, whereas all trucks coming from Mexico through FAF2 region San Diego and going to a destination further east choose Tecate in the simulation (see Figure 6 for a map with border crossings).

The model would only be able to route trucks into San Diego using Tecate if congestion around Otay Mesa would delay travel times significantly. This simplification has been made to account for limited information about origins and destinations across the border in Mexico, and the user should have in mind that this limits the ability of the model to simulate scenarios affecting border crossing times.

A major simplification limiting scenario testing is the absence of a mode choice model. Though the FAF2 data provide commodity flows by mode, the external model only processes flow by truck. The internal model generates truck trips only and ignores other modes, such as rail or water. Therefore, no mode shift can be represented in the model. Even if California extended the rail system substantially, making rail faster, more reliable and less expensive, the SANDAG truck model would continue to generate the same number of trucks.

Similarly, the time of day is a fixed share. If SANDAG implemented a policy where tolls at peak hours are higher than off peak, the truck model would try to shift trucks to other routes avoiding the toll roads at peak hour, however, no shift in time of day would be modeled. The effect most likely found in reality, that more trucks try to schedule their time of departure to reduce toll costs, could only be modeled by exogenously adjusting the time of day trucks are generated. In general, it is expected that long-distance truck trips are less sensitive to tolling. This effect barely can be modeled, as the information from the FAF2 data about the final destination of a truck trip is lost in the processing once the truck leaves San Diego County.

Change in transportation costs, either through tolls or increased fuel prices, is likely to change consolidation of truck loads and fleet composition in the long run. With higher transportation costs empty loads become much more expensive, and trucking companies will try to better consolidate truck loads to reach a higher degree of capacity utilization. In the model, changing transportation costs should impact truck production and attraction rates, as well as the trip distribution friction. In reality, economic activity also has a major impact on truck traffic. In times of recession, consumption declines, if the recession is global exports are reduced at the same time. The current implementation, however, works with static trip rates and friction factors that need to be changed exogenously if necessary.
Knowledge about the military sites is limited, if in existence at all. Employment statistics report more than 85,000 military employees in San Diego County in 2004. The military sites generate a large amount of truck traffic, both for food and general living supply as well as military material supply. No specific truck counts for military sites were available. The assumptions made in the special generator module (section 5.2) are very crude and call for refinement. In this current implementation, it is not recommended to use to model to test policies that specifically aims at changing military truck traffic.

A major concern remains the highly limited data on truck traffic for model calibration and validation. Only the WIM counts are considered to be a reliable data source for validation, as explained in section 8. The WIM counts, however, are all located on major highway roads. In consequence, it is not possible to prove that the model is working properly on minor roads, other than that the overall picture of truck flows appear to look reasonable. Special care should be taken if the model is applied in small area studies that focus on non-highway roads. While the model captures the main trend, such as having fewer trucks on roads with less capacity, higher congestion or lower speeds, the absolute number on one link of a smaller road should be evaluated carefully before being used for major policy decisions. There is high confidence that the direction of change is simulated correctly, such as a reduced number of trucks on a road where capacity has been reduced. The amplitude of this change, however, should be taken with care. Finally, data on trip length distribution would be desirable for model validation.

10.2 Recommendations for future truck counts

The model development revealed that a major bottleneck for a thorough validation is the lack of reliable count data (compare section 8). Most truck counts were done as a one-point-in-time count. Therefore, any special circumstance on that particular day, such as an accident, a construction site or a sports event, is reflected in the count data. The model, in contrast, aims at simulating an annual average weekday (AAWD). WIM data has shown that the standard deviation at an average count station is as high as 15 percent, and single days can be off the AAWD by several dozen percent. The only available data where the same site was counted for a larger number of days are the WIM counts. Unfortunately, there were only 10 WIM stations, with seven station having bi-directional counts, providing a total of 17 count sites. Though the quality of the WIM counts is very high, the number of WIM sites is very low for a comprehensive model validation. Therefore, it is a strong recommendation that SANDAG should improve truck count data by collecting truck data over several days at a larger number of major roads.

To identify locations that are most desirable to validate a truck model, a synthetic assignment has been set up. In this synthetic assignment, a trip table with 0.01 truck trips is assigned from every zone to every zone. With 4,593 zones plus 12 external zones in San Diego County, a total of almost 200,000 truck trips were assigned to the highway network. With no personal vehicles in this assignment, every truck will use the shortest route from the origin to the destination zone. The number of trucks on any given link is an indicator of how many zones are connected by this link. The simulated volumes on
each link in this synthetic assignment are shown on the ordinate in Figure 20. The absissa shows the volume on these links in a regular truck flow assignment. Dots that are found in the upper right corner are considered to be the most important points, as they both connect many zones and receive high volumes in a truck assignment. The red lines show dots of links that are larger than the average in either dimension. By definition, the number of dots in the upper right quadrant of the red lines is one quarter of all links. Hence, this method helps to easily eliminate 75 percent of all links as non high-priority links for truck counts. Since 25 percent of all links still is a high number of sites to count, the green lines show the median of all counts. Due to a few fairly large volumes, the median is larger than the average and, therefore, reduces the number of most important count locations.

Figure 20: Comparison of synthetic and actual assignment to identify important links
Figure 22: Significant links with volumes above the median

Figure 21: Significant links with volumes above the average
Figure 22 and Figure 21 show maps where the links highlighted that fall into this upper right quadrant in Figure 20, one for the median and one for the average, respectively. These two plots give some idea which sites might be the most important ones to count trucks for future model validation. Obviously, the number of links that could be considered for truck counts remains large and probably exceeds a wise spending on resources. Local knowledge should be use to finally select site for truck counts among these links that have been identified as being significant for a truck model development.
References


### Appendix A: Input file for special generators

File `<specialGenerators.csv>`

<table>
<thead>
<tr>
<th>TAZ</th>
<th>Zone Description</th>
<th>Data in five year intervals</th>
<th>Truck generation factor</th>
<th>Truck shares</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>MilitaryEmployees</td>
<td>Y2000 Y2005 Y2010 ... Y2050</td>
<td>trkProd trkAttr lhdShare mhdShare hhdShare</td>
<td></td>
</tr>
<tr>
<td>58</td>
<td>MilitaryEmployees</td>
<td>21 21 21 21</td>
<td>0.02 0.02</td>
<td>0.307 0.155 0.538</td>
</tr>
<tr>
<td>1979</td>
<td>MilitaryEmployees</td>
<td>28414 28414 28414 28414</td>
<td>0.02 0.02</td>
<td>0.307 0.155 0.538</td>
</tr>
<tr>
<td>3370</td>
<td>MilitaryEmployees</td>
<td>9325 9325 9325 9325</td>
<td>0.02 0.02</td>
<td>0.307 0.155 0.538</td>
</tr>
<tr>
<td>3484</td>
<td>MilitaryEmployees</td>
<td>4354 4354 4354 4354</td>
<td>0.02 0.02</td>
<td>0.307 0.155 0.538</td>
</tr>
<tr>
<td>3509</td>
<td>MilitaryEmployees</td>
<td>2456 2456 2456 2456</td>
<td>0.02 0.02</td>
<td>0.307 0.155 0.538</td>
</tr>
<tr>
<td>3542</td>
<td>MilitaryEmployees</td>
<td>88 88 88 88</td>
<td>0.02 0.02</td>
<td>0.307 0.155 0.538</td>
</tr>
<tr>
<td>3733</td>
<td>MilitaryEmployees</td>
<td>1786 1786 1786 1786</td>
<td>0.02 0.02</td>
<td>0.307 0.155 0.538</td>
</tr>
<tr>
<td>3734</td>
<td>MilitaryEmployees</td>
<td>14395 14395 14395 14395</td>
<td>0.02 0.02</td>
<td>0.307 0.155 0.538</td>
</tr>
<tr>
<td>3890</td>
<td>MilitaryEmployees</td>
<td>569 569 569 569</td>
<td>0.02 0.02</td>
<td>0.307 0.155 0.538</td>
</tr>
<tr>
<td>4035</td>
<td>MilitaryEmployees</td>
<td>712 712 712 712</td>
<td>0.02 0.02</td>
<td>0.307 0.155 0.538</td>
</tr>
<tr>
<td>4068</td>
<td>MilitaryEmployees</td>
<td>10493 10493 10493 10493</td>
<td>0.02 0.02</td>
<td>0.307 0.155 0.538</td>
</tr>
<tr>
<td>4069</td>
<td>MilitaryEmployees</td>
<td>5076 5076 5076 5076</td>
<td>0.02 0.02</td>
<td>0.307 0.155 0.538</td>
</tr>
<tr>
<td>4513</td>
<td>MilitaryEmployees</td>
<td>7213 7213 7213 7213</td>
<td>0.02 0.02</td>
<td>0.307 0.155 0.538</td>
</tr>
<tr>
<td>3450</td>
<td>DailyAirMailTons</td>
<td>168 168 168 168</td>
<td>0.02 0.02</td>
<td>0.307 0.155 0.538</td>
</tr>
<tr>
<td>3641</td>
<td>DailyCruiseShips</td>
<td>480 480 480 480</td>
<td>0.1 0.1</td>
<td>0.307 0.155 0.538</td>
</tr>
</tbody>
</table>

### Appendix B: Manual to run the external model

To run the external model, three steps are required. First, a java program runs to analyze the FAF2 data and create truck trip tables. Second, a temporal assignment needs to be done in TransCAD to calculate the number of trucks entering and leaving San Diego County at every external station. And finally, a Fortran program needs to be run to
reformat the TransCAD output of the external model into TransCAD input of the internal model.

**Step 1: Java program**

Java Runtime Environment (JRE) needs to be installed. The model has been developed using Java(TM) 6 Update 7 (Version 1.6.0.70). The external model is run from a command window:

```
<ext_model_path>/sandagExternal.jar <ext_model_path>/javaFiles/sandag.properties
```

The model analyzes FAF2 flows for the years 2002, 2010, 2015, 2020, 2025, 2030 and 2035. The model further adjusts the forecast for the years 2010 to 2035 and extrapolates future forecasts for the years 2040 to 2050 in five-year increments.

The directory `ext_model_path>/input/freight` has a file `ForecastNoOfTrucksInUS.csv` that is used to adjust the FAF2 forecast. This file contains the following data:

<table>
<thead>
<tr>
<th>Year</th>
<th>modeledNoOfTrucks</th>
<th>FAFforecastNoOfTrucks</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>2837199.453</td>
<td>2837199.453</td>
</tr>
<tr>
<td>2015</td>
<td>3125001.03</td>
<td>3125001.03</td>
</tr>
<tr>
<td>2020</td>
<td>3469247.002</td>
<td>3469247.002</td>
</tr>
<tr>
<td>2025</td>
<td>3761219.265</td>
<td>3904438.04</td>
</tr>
<tr>
<td>2030</td>
<td>4053191.528</td>
<td>4453044.421</td>
</tr>
<tr>
<td>2035</td>
<td>4345163.791</td>
<td>5140217.022</td>
</tr>
<tr>
<td>2040</td>
<td>4637136.054</td>
<td>-1</td>
</tr>
<tr>
<td>2045</td>
<td>4929108.317</td>
<td>-1</td>
</tr>
<tr>
<td>2050</td>
<td>5221080.58</td>
<td>-1</td>
</tr>
</tbody>
</table>

The first column contains the years that may be adjusted. The second column `modeledNoOfTrucks` defines the total number of trucks per day in the entire U.S. Values in this column can be adjusted by the user to specify a different growth in truck trips than proposed in section 6.2. The third column `FAFforecastNoOfTrucks` never should be altered. This column is meant to be a reference point for the user to understand what the FAF2 forecast proposes and, thus, how far the forecast has been changed by the user. The model uses the third column to determine if an adjustment is necessary or not. In the shown example, the values for the years 2010 to 2020 are identical in the second and third column, telling the model that no adjustment is needed. If the values in columns two and three are different, the value in column three is irrelevant for the model; the simulated number of truck trips is scaled to the number of trucks defined in column two.

**Step 2: TransCAD assignment**

The java code produces truck trip tables for every simulated year across the entire U.S. A temporal assignment is needed to define how many trips enter and leave San Diego County at any given external station. The TransCAD file `RegTruckModelMap.map`
SANDAG Truck Model

provides a basis map to get started. The output file of the java code `<ext_model_path>\base\dailyCountyTruckFlows<year>.csv` needs to be opened in TransCAD and is reformatted to a matrix using the command Matrix -> Import… to create an OD matrix (RowID = FromID, ColumnID = ToID, Field = Trucks). An index has to be set to reference the county FIPS to the node ID in the TransCAD network. The file `nodesOfCountyFips.csv` contains this reference. The function Matrix -> Indices… -> Add Index… (Dataview = nodesOfCountyFips, Field (top) = FIPS, Use for = Both, Field (bottom) = Node) allows to set the node IDs as row and column names. For a subsequent subarea analysis, the user has to select the FAF zone San Diego called with full name "San Diego-Carlsbad-San Marcos, CA MeSA" and abbreviated as "CA San D". Next, set the highway layer as the working layer. Using Planning -> Planning Utilities -> Subarea Analysis… the user can create the subarea and specify fields for assignment (Network = geography\HighwayNetwork.net, Method = User Equilibrium, Time = [Travel time in min], Capacity = Daily Capacity, Preload = BackgroundVolume). Figure 23 shows the settings in TransCAD. After completion of the subarea analysis TransCAD writes out two files, one OD-Matrix for the selected subarea and one binary file with link volumes. The matrix file is used in step 3 and needs to be exported into a csv file. Using the function Matrix -> Export… allows to export the matrix to a table where one record is written for each cell. A csv file has to be exported for step 3. It is recommended to add the year to the file name.

Step 3: Fortran program

Finally, the csv file exported in step 2 is reformatted by a Fortran program. This Fortran program can be called in two different ways. The simple version only reformats the TransCAD output file (recommended for simulation years other than 2002 and 2030):

`<ext_model_path>/reformat.exe <reg_model file>`

This program reads the csv file generated in step 2 and splits flows into IE, EI and EE trips. These flows are written to three csv files. These three files need to be copied into the following directory of the internal truck model:

`<int_model_path>\tc\res\regModel\regionalIEtrips.csv`
`<int_model_path>\tc\res\regModel\regionalEItrips.csv`
`<int_model_path>\tc\res\regModel\regionalEEtrips.csv`

The program is also prepared to update the `cordon.prn` file. This file contains EI and IE trips of all vehicles. With this truck model trucks are simulated explicitly and, therefore, need to be subtracted from the cordon.prn file. The `cordon.prn` file carries two years, 2003 and 2030. To update the cordon file, the regional model result file need to be
available for the years 2002 (closest available to 2003) and 2030. To update the cordon.prn file, the reformat program is called as follows:

<ext_model_path>/reformat.exe <reg_model file 2002> <reg_model file 2030> <cordon file>

To avoid confusion whether a cordon file has been updated already, the updated data is saved into a new file called cordon_noTrucks.prn. This file needs to be copied to the corresponding directory of the local travel demand model.

The connection between the external and the internal model is complete. The next time the GISDK script of the internal model is executed, the external flows are added to the SANDAG truck model. To run the local truck model, the path in the script TruckModel.rsc in line 8 needs to be adjusted to the local settings. The model is started with the Macro name of the add-in run.