A Working Demonstration of a Mesoscale Freight Model for the Chicago Region

Final Report and User's Guide

prepared for the

Chicago Metropolitan Agency for Planning

prepared by

Cambridge Systematics, Inc.

with contributions from

Kouros Mohammadian
Kazuya Kawamura
Joshua Auld

www.camsys.com

Final Report and User's Guide

A Working Demonstration of a Mesoscale Freight Model for the Chicago Region

prepared for the

Chicago Metropolitan Agency for Planning

prepared by

Cambridge Systematics, Inc.
115 South LaSalle Street, Suite 2200
Chicago, IL 60603

date

June 30, 2011

Table of Contents

Table of Contents

(continued)

List of Tables

List of Figures

Executive Summary

This report describes the development and implementation of a powerful and innovative prototype model of freight movements that has been prepared for the Chicago Metropolitan Agency for Planning (CMAP). The prototype model provides a framework for analyzing a variety of important goods movement decisions that are made by individual businesses. The objective of the prototype model at this early development stage is to provide CMAP with a working demonstration of a theoretically robust framework.

The prototype model that is documented in this report – also called the mesoscale model – is anticipated to serve as the middle layer of a three-layered analytical framework. The mesoscale model will bridge the proposed macroscale and microscale models as follows:

- The proposed macroscale model will use economic models to generate high-level commodity flow data that are similar to the Federal Highway Authority's Freight Analysis Framework (FAF3) data. These data are anticipated to be further evaluated in the mesoscale model.
- The mesoscale model effectively breaks down the high-level commodity flows into a freight trip table that uses zone sizes which are suitable for regional-level analysis.
- The proposed microscale model, which will use outputs from the mesoscale model, will provide a way to examine detailed freight vehicle movements.

This innovative multilayered framework of analytical tools stems in part from earlier proposed frameworks that are described in Section 3.0.

■ Model Overview

Section 1.0 describes the model framework. This section describes the types of goods movements and decision makers that are modeled; the types of decisions that are modeled; and other model details such as the zone system and network.

The analytical power of the mesoscale model is attributable in part to its focus on modeling decisions at a detailed level. Unlike conventional freight models, the mesoscale model was designed with the capability of accommodating a high level of detail in every step and is multimodal in its approach. State of the practice freight travel demand models normally perform highly aggregated analyses, focusing on travel behavior at the zonal level, and generate trips only for the truck mode.

The mesoscale model **analyzes goods movement decisions that are made by individual businesses**, which also are referred to as the model agents. By focusing on individual businesses, the model can incorporate a broad range of economic and other business strategies that drive the freight-related decisions of businesses.

In addition, the mesoscale model **evaluates transportation and logistics paths** at a detailed level. Path-related decisions such as what combination of modes to use, how frequently to make shipments, and whether or not to use a logistics handling facility (such as a distribution center) are modeled. In other words, the model adopts a holistic perspective of path selection that incorporates tradeoffs between transportation-related factors, such as mode choice, and factors that are not specifically transportation-related, such as inventory costs.

The mesoscale model is **multimodal** and includes truck, rail (including carload and intermodal), air, and water in the path selection process. The model uses an EMME/3 network-based process to generate paths for each of these modes.

While the model is capable of incorporating substantial detail, the model framework is also very flexible and can be tailored to use less detailed input data or to produce output with less detail. Analysis with the mesoscale model can be less detailed if the necessary data are not available. Furthermore, the model can be customized to produce a variety of output data to suit different analysis needs.

■ Model Components

Section 2.0 discusses the data needs of the model. Data inputs for the prototype model and anticipated data needs for the full implementation of a regional freight model are described.

Section 3.0 provides a step-by-step description of the model stream. The mesoscale model evaluates freight movements as follows. First, in ***Firm Generation***, individual firms in the U.S. and abroad that produce and/or consume commodities are synthesized. Second, in ***Supplier Selection***, individual firms are paired together. The resulting supply chains generate the physical transport of commodities between suppliers and buyers. Third, during ***Flow Apportionment***, high-level commodity flows between regions are transformed into total annual shipment volumes between suppliers and buyers. Finally, in ***Path Selection***, the selection of transport and logistics paths that are used for transporting individual shipments from supplier to buyer is modeled.

■ Application

Section 4.0 describes tests that were conducted during the prototype development to ensure that the basic functions of the model are working as expected. Section 4.0 also describes potential applications of the model. Following an anticipated data collection and calibration effort, the model is expected to be used by the Chicago Metropolitan Agency for Planning to address a variety of freight-related questions that cannot be adequately addressed by conventional freight modeling tools. The first three sections of this report describe the capabilities of the mesoscale model. The last section discusses the relevance of the mesoscale model to the analysis of specific infrastructure, policy, and operational issues.

Finally, Section 5.0 comprises a User's Guide with instructions on setting up and using the model.

■ Limitations

While this model development effort has resulted in a very promising prototype, it is still just a prototype and it may be premature to be considered for evaluation purposes for the following reasons:

- First, because the macroscale model has not yet been developed, through trips (External-to-External) are not included in the model. This is an important source of freight traffic in the Chicago region;
- Second, a more detailed regional transport and logistics network is needed to understand logistics-related travel throughout the regional network;
- Third, the choice models and much of the data use placeholder values. Since the model framework is very thorough, it will need and benefit from an equally thorough calibration and validation process.

■ Next Steps

The limitations of the prototype that are described above will be resolved through data collection and further model development.

The mesoscale model contains many features that will require calibration and validation using observed data. The following types of data will be needed:

- Enhanced network data for all freight modes in the region;
- Supplemental base-year business data (construction, agriculture, and foreign employment) and future-year business data;
- Path-related data for the calibration of level-of-service and path cost parameters; and

- Firm surveys to better understand and calibrate the **Supplier Selection** and **Path Selection** models.

Furthermore, the macroscale and microscale models are expected to be developed and implemented at some point in the future. The macroscale model is especially important for the mesoscale model because its primary output – high-level flow data – comprises one of the main inputs to the mesoscale model. In addition, through trips in the mesoscale model are expected to be provided by the macroscale model.

Following the data collection effort, an intensive model calibration and validation process will be required to develop, test and validate the mesoscale model to reflect freight flows in the Chicago region.

1.0 Framework

The prototype mesoscale freight model provides a powerful and innovative framework for analyzing freight movements to, from, and within the Chicago region. Following an anticipated data collection and calibration effort, the model is expected to be used by the Chicago Metropolitan Agency for Planning to address a variety of freight-related questions that cannot be adequately addressed by conventional modeling tools, which typically evaluate travel behavior at an aggregate level and generate trips only for the truck mode. For example, a typical freight travel demand model may generate truck trips based on total employment in broad industry sectors and distribute truck trips between zones based on travel time. In contrast, the mesoscale framework:

- Models the goods movement decisions of the individual business;
- Uses detailed North American Industry Classification System (NAICS) industry classes to inform decision-making in the modeling process;
- Evaluates transportation decisions based on numerous factors, including travel time, distance, mode, and inventory costs; and
- Includes truck, rail, water, and air as modes.

Furthermore, the prototype mesoscale model focuses on modeling commodity supply chains. The prototype model is designed to transform high-level commodity flows from a macroscale model into individual shipments between individual shippers and receivers and to evaluate transport and logistics choices, such as the mode decision, at this highly disaggregate level. The primary macroscale input at this time is the FAF3 dataset, which contains aggregate data on commodity flows between broad geographic regions with unspecified shippers and receivers, paths, and intermediate stops.

The mesoscale model performs four critical processes to transform the FAF data into commodity tours at a suitable level of detail. These four steps accomplish the following objectives:

1. In ***Firm Generation***, individual firms that produce and/or consume commodities the U.S. and abroad are synthesized;
2. In ***Supplier Selection***, individual firms are paired together to form supply chains that represent the physical transport of commodities between suppliers and buyers;
3. During ***Flow Apportionment***, high-level commodity flows between regions are transformed into total annual shipment volumes between suppliers and buyers; and

4. The ***Path Selection*** stage is used to model individual shipments that are passed from supplier to buyer, and to model the transport and logistics paths that are used by the shipper-receiver team for the purpose of transporting the modeled shipments.

Ultimately, the mesoscale model is anticipated to serve as a bridge between the proposed macroscale and microscale models. The proposed macroscale model will use high-level economic modeling tools to generate high-level commodity flow data that are similar to the FAF3 data. These data are anticipated to be further evaluated in the mesoscale model. The mesoscale model effectively breaks down the high-level commodity flows into a freight trip table that uses zones which are sized for regional-level analysis. The proposed microscale model, which will use outputs from the mesoscale model, will provide a way to examine detailed freight vehicle microsimulations. This innovative multilayered framework of analytical tools stems in part from earlier frameworks (described in Section 3.0).

Due to the innovative nature of the prototype mesoscale model, a significant amount of discourse was undertaken during the framework development to identify the most meaningful and promising ideas for implementation. These ideas fulfill the following objectives:

- Meet CMAP's stated need to model freight vehicles on the CMAP regional travel demand model network;
- Consider the underlying economic perspectives of the agent(s), as appropriate;
- Have a history of academic development or acceptance;
- Implement using readily available software; and
- Be able to calibrate using data that CMAP can realistically obtain.

The remainder of this section discusses the elements of the mesoscale model in general terms. First, the types of freight movements and agents that are modeled are described. The conceptual focus on certain movements and agents has implications for the practical implementation of the model components. This section describes these implications. Second, the network, zone system, and modes that were developed to support the modeling of these freight movements and agents are described. Third, the agent-based evaluation of transport and logistics costs and decisions that has been implemented in the model is presented.

1.1 Freight Movements and Their Agents

This section describes the type of freight movements that will be modeled and the types of decision-making agents that are generated to support the objective of modeling goods movements. The mesoscale model focuses on movements of commodities and the types of businesses that produce or use commodities. The concepts that are described in this section provide the framework for the steps that are carried out during **Firm Generation**, **Supplier Selection**, and **Flow Apportionment**.

1.1.1 Commodities Versus Non-Commodities

The mesoscale model focuses on modeling the movements of commodities, which are defined as identifiable goods that have value. The main prototype input data source – the Freight Analysis Framework, or FAF – summarizes flows of commodities throughout the U.S. based on the commodity classes that are available in the Commodity Flow Survey (CFS). Because FAF3 is considered to be most accurate for these types of freight movements, the mesoscale model will focus on the commodities that are reported in FAF3. FAF3 reports commodity classes using the two-digit Standard Classification of Transported Goods (SCTG) codes.

The mesoscale model does not model movements of unidentified commodities, which comprises a commodity category in the FAF3 data, or service vehicles. Furthermore, because the mesoscale model focuses on commodity moves, it cannot be used to understand and address non-commodity commercial vehicle movements (including trips with service, utility, construction, and maintenance purposes).

The mesoscale prototype attempts to model flows of waste/scrap products, which are included in the latest FAF3 release. However, the prototype model may not provide a meaningful fit for this category due to categorical mismatches and lack of information on the producers and consumers of waste and scrap. The categorical mismatch occurs when the movements of scrap metal products that are produced by select metal working industries are modeled, whereas waste products (such as recyclables) are not specifically modeled. However, flows from the entire waste/scrap category are apportioned to these select scrap metal producers and their buyers. This likely has undesirable consequences such as the overestimation of scrap metal flows. Furthermore, the framework only identifies two NAICS industries as producers of scrap metal, when in reality there may be additional industries in this category. This area should be revisited in the full mesoscale implementation.

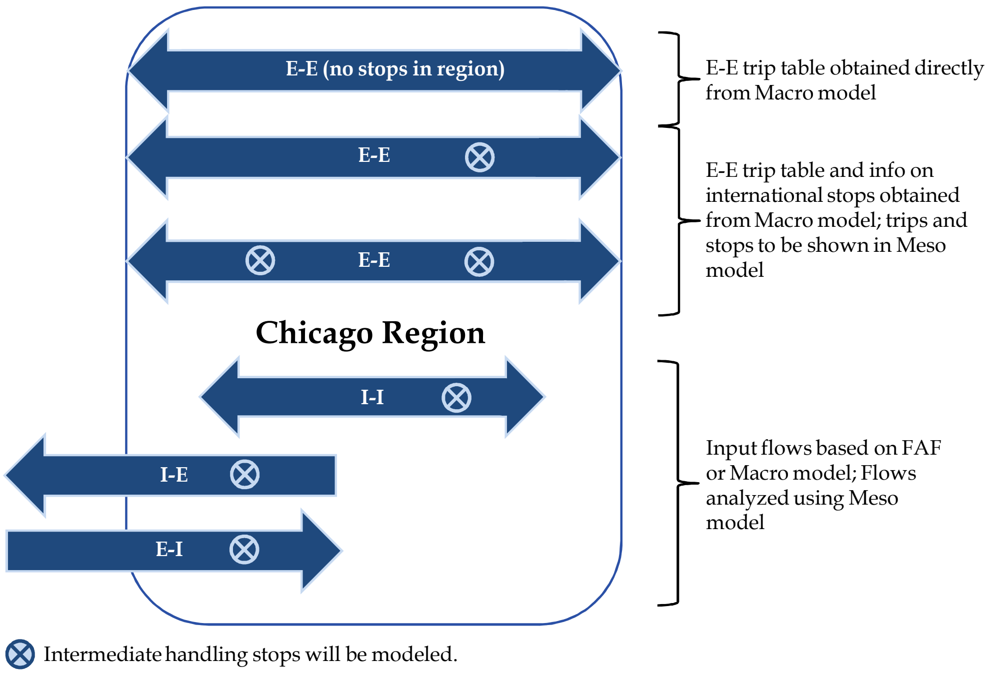
The focus on commodities has the following implications for the mesoscale model. First, the agents in the mesoscale are firms that either make or use one or more SCTG commodities. Second, the mesoscale model links together supplier firms that generate a particular commodity with buyers that use or consume the same commodity. Third, high-level commodity flows from FAF3 are apportioned to individual firms during the modeling process.

1.1.2 Categories of Freight Moves

The mesoscale model focuses on processing commodity moves as reported in the FAF3 data, which primarily consist of long-haul external-to-internal and internal-to-external (E-I/I-E) movements. Some I-I movements also are included. For example, petroleum delivered by truck from the regional BP refinery to local gas stations may be represented.

Freight movements through the region – external-to-external (E-E moves) – are expected to be provided by the macroscale model[1] (Figure 1.1). Through movements that undergo no intermediate handling in the Chicago region can be input directly alongside vehicles from the mesoscale model "as is" for assignment. Through movements from the macroscale model can also consist of multi-modal movements with one or more stops in the region. For example, most rail and water moves that travel through the region probably undergo intermediate handling by trucks. These secondary truck movements for E-E flows are expected to be provided by the macroscale model.

Figure 1.1 Mesoscale Model Inputs



Local pick-up and delivery traffic to some extent will be included but not in a deliberate way. For example, the I-I commodity move of petroleum from a regional oil refinery to a local gas station also could be classified incidentally as local delivery traffic. Other than these incidental movements, local pick-up and delivery traffic will not be represented.

[1] Information on E-E flows cannot be obtained from the FAF3 except by inference of shortest paths which may pass through the Chicago region. The FAF3 provides flow information from the ultimate origin to the ultimate destination without any information of the number or geographic location of any intermediate stops that may be involved in the shipment of that commodity. While intermediate handling information is reported in the STB Waybill database which is used by the FAF to develop flows by the rail mode, that information is discarded when that rail information is reported in the FAF3.

The use of macroscale commodity flow data necessitated a meaningful representation of long-haul freight movements in the mesoscale framework. This led to the generation of agents in the entire U.S., the development of a national-level network and zone system, and the development of foreign agents and distance data. This provides a coherent and transparent environment for apportioning the high-level commodity flows to individual agents.

1.1.3 Agents and Supply Chains

Most long-haul commodity moves are coordinated between three decision-makers: shippers, receivers, and carriers or third-party logistics (3PL) firms. The key decision-maker is the receiver (the business that is receiving the shipment). The receiver specifies the critical parameters that must be met such as mode, cost, reliability, and delivery time. The shipper is responsible for meeting these requirements.

As a result, the shipper-receiver pair – also referred to as a supply chain – essentially functions as one decision-making unit. This is clearly the case when the shipper and receiver belong to the same company. If they are from different companies, then their objective functions are slightly different (e.g., the shipper's objective function includes a profit component). However, in both cases, the shipper and receiver seek to minimize their total transport and logistics costs.

For-hire carriers or 3PL firms that handle freight moves for the shipper-receiver pair are also under obligation to fulfill the requirements set forth by the receiver. The carrier typically has discretion over other decisions such as route and intermediate handling choices as long as these decisions remain consistent with the overall control variables. For example, there may be instances where a carrier deviates from the most direct route in order to pick up a driver or to pick up an additional partial load,[2] thus incurring extra Vehicle-Miles Traveled (VMT) and an intermediate handling stop that are not necessary from the perspective of the shipper-receiver pair. These relatively spurious moves are not modeled as part of the prototype model. We recommend ignoring these moves in the full mesoscale implementation because these moves probably make a nominal contribution to VMT, have significant data requirements,[3] and would likely pose major issues in validation.

[2] I.e., a carrier may find it worthwhile to carry the occasional load from a nearby plant without having enough business in the area to make it worth building a distribution center. In the view of the modeled shipper or receiver, this route would not be used because it would incur extraneous VMT. However, the variability of operations among carriers and within the operations of a single carrier fleet makes these movements difficult to represent accurately.

[3] Obtaining the data that are required to model the decisions of both the shipper-receiver pair and the carrier or 3PL has substantial practical challenges. When a carrier facilitates a move between the shipper and receiver, the individual businesses are typically unconcerned and/or unaware of each others' transportation decisions. This creates difficulties in finding the right individuals to survey, getting them to participate, and generating realistic choice sets for each survey.

In addition to shipper and receiver firms that make or use commodities, wholesale firms are represented in the mesoscale model. Although wholesale firms are behaviorally closer to third-party firms than to manufacturers or consumers of a product, wholesale businesses were a significant part of the Commodity Flow Survey (CFS) sampling framework. As a result, the FAF3 flow data (which are based primarily on CFS data) contain a significant number of shipments that involve a wholesale business. Because of this, representation of wholesale firms in the mesoscale model was considered to be a valuable addition to the framework.

In summary, the shipper-receiver pair is the decision-making unit in the mesoscale freight model. Wholesale firms also are represented. The implications of this focus for the mesoscale model are as follows. First, shippers (or suppliers) that produce a particular good are paired with receivers (or buyers) who use the same good. Second, in some cases, the supply chain link that is formed is actually a shipper-to-wholesaler or a wholesaler-to-buyer chain. Third, the resulting shipper-receiver pairs are the agents to which the commodity flows are apportioned.

The resulting behavioral framework fulfills important elements of the meaningful and promising modeling approach that was described earlier. Most importantly, the economic decisions of shipper-receiver pair drive the shipping process. Furthermore, this representation has a history of academic acceptance (e.g., in the Freight Activity Microsimulation Estimator, or FAME, and as described later in Section 3.0), is implementable using readily available software, and has a realistic chance of successful calibration using survey-based data.

■ 1.2 Network and Mode Overview

This section describes the network and modes that are implemented in the mesoscale model. These features were developed to support the global reach and multimodal provisions of the mesoscale framework.

1.2.1 Zone System and Network

The mesoscale zone system is comprised of township-sized zones in the inner CMAP counties, county-sized zones on the fringes of the CMAP region, and FAF3 zones elsewhere. For the prototype model, a rudimentary national ground transportation network was developed along with corresponding generic Class I rail routes. A limited water network was developed in the Great Lakes area. A relatively detailed network of truck routes was developed within the CMAP region.

Logistics nodes that represent logistics handling facilities also were developed for the prototype. These cover activities such as break-bulk handling, intermodal lifts, transloading and distribution/consolidation.

1.2.2 Modes

The mesoscale model focuses in greatest detail on modeling truck, rail, and truck-rail intermodal modes. The modes are further distinguished by type of carriage: less-than-truckload (53'), truckload (53'), truck with container (40'), carload, and intermodal (single-stack, 40' containers).

Water and air modes are modeled but with less network detail. A rudimentary water network was developed for Great Lakes traffic while air distances are simply assumed to be the same as ground distances.

For the prototype, water and air moves are assumed to travel through logistics nodes in the Chicago region and in the external region. For example, it is assumed that air and water cannot provide direct movement between the supplier and buyer without drayage to a water port or airport. If desired, this constraint can be relaxed by coding water network access links directly to supplier or buyer sites.

1.3 Path-Based Analysis

Path-based analysis is a fundamental component of the mesoscale model. Decisions regarding mode, shipment size, inventory considerations, and other logistics concerns are handled using this methodology.

First, a set of feasible paths between each O-D pair is enumerated. The EMME/3 mesoscale network provides the data for most of the path-building. The model network data are supplemented by additional data to cover paths between the Chicago region and Alaska, Hawaii and foreign countries.

Second, a plausible set of parameters is applied to the EMME/3 path skims to generate total annual transport and logistics costs for each combination of path and shipment size. Four different shipment sizes are evaluated for the prototype. The utility of the entire path is modeled.

The literature was explored to identify a suitable formulation of transport and logistics costs for this model. This review is described in the following section.

1.3.1 Background: Models from Other Areas

Advanced freight models that have been implemented thus far in the U.S. primarily have focused on truck tours. Truck vehicle touring models have been developed in Ohio[4] and Calgary;[5] however, neither effort models the logistics handling of commodity flows. Furthermore, existing truck touring models essentially represent the perspective of carriers (i.e., they analyze the demand for trucks) rather than the underlying economic need for the commodity itself (as the mesoscale model does). The Oregon statewide model[6] has a commercial transport component that addresses logistics handling, but it simply uses a random sampling of observed data to replicate observed outcomes rather than explanatory models with a behavioral basis.

In contrast, the prototype mesoscale model approaches commercial vehicle movements from the perspective of individual businesses that produce or consume goods. The genesis of this approach can be traced to a handful of innovative research efforts that focus on framework development or implementation and testing.

One important predecessor to the macroscale-mesoscale-microscale framework is the SMILE[7] transport and logistics model developed in the Netherlands. The SMILE model simulates the goods movement cycle in three steps: goods production and distribution, warehouse location and inventory chains, and multimodal assignment to the network.

Another important framework was developed by Fischer, et al. for Los Angeles County Metropolitan Transportation Authority.[8] This three-layered approach calls for the modeling of economic trade relationships between freight producers and consumers; identifying logistics decisions that are made in transporting the goods between producers and consumers; and assigning the resulting vehicles to a transportation network.

---

[4] Gliebe, J. P., O. Cohen, and J. D. Hunt. A Dynamic Choice Model of Commercial Vehicle Activity Patterns. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2003. TRB, National Research Council, Washington, D.C., 2007: 17-26.

[5] Hunt J. D. and K. J. Stefan, 2007, Tour-based microsimulation of urban commercial movements. *Transportation Research 41B*:981-1013.

[6] Hunt, J. D., R. R. Donnelly, J. E. Abraham, C. Batten, J. Freedman, J. Hicks, P. J. Costinett, and W. J. Upton. Design of a Statewide Land Use Transport Interaction Model for Oregon. *Proc., 9th World Conference for Transport Research*, Seoul, South Korea, July 2001.

[7] Tavasszy, L.A., M.J.M. van der Vlist, C.J. Ruijgrok, and J. van der Rest. Scenario-wise analysis of transport and logistics systems with a SMILE. Accessed December 29, 2010 at: http://www.tongji.edu.cn/~yangdy/news/_PRIVATE/softw1.htm

[8] Fischer, M., Outwater, M., Cheng, L., Ahanotu, D. and R. Calix. An Innovative Framework for Modeling Freight Transportation in Los Angeles County. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1906. TRB, National Research Council, Washington, D.C., 2005: 105-112.

Additionally, the Aggregate-Disaggregate-Aggregate (ADA) framework that has been proposed for and partially implemented in Norway and Sweden[9,10] is relevant to the proposed mesoscale framework. The following ADA modeling steps are very similar to the steps that were developed for the mesoscale model:

1. Disaggregation of commodity flows at their production and consumption ends to firm-to-firm flows, where shipping and receiving firms are paired and are then treated as a single behavioral unit;

2. Modeling of logistics decisions that are made by the shipper-receiver pair based on evaluation of the total transport and logistics costs on available paths and under various scenarios, such as different shipment frequencies; and

3. Aggregation of individual shipments to origin and destination zones for network assignment purposes.

A modern and comprehensive accounting of transport and logistics costs also was developed for the ADA framework as part of Step 2. This formulation, which is described further in Section 3.2, is used in the mesoscale model.

The mesoscale model effectively implements Steps 1 and 2 of the ADA framework. Step 3 also is implemented in the prototype mesoscale model primarily as a means of checking the reasonableness of the results.

Finally, the Freight Activity Microsimulation Estimator[11] (FAME) software that was developed at the University of Illinois-Chicago has implemented elements of the above frameworks using readily available data. While logistics decision modeling in FAME is relatively simplified, the basic components and inputs of this development have been critical to the development of the mesoscale model.

1.3.2 Transport and Logistics Cost Formulation in the Mesoscale Model

This section describes the transport and logistics cost formulation that is used in the mesoscale model. The ideal formulation models logistics decisions in a joint fashion by capturing all transport and logistics costs in a single equation. This reflects real-world decision-making of freight movers more accurately than a series of choice models would.

---

9 de Jong, G. and M. Ben-Akiva, A micro-simulation model of shipment size and transport chain choice, Transportation Research Part B: Methodological, Volume 41, Issue 9, November 2007, Pages 950-965.

10 Ben-Akiva, M. and G. De Jong (2008). The Aggregate-Disaggregate-Aggregate (ADA) Freight Model System, In: Recent Developments in Transport Modeling. Emerald, 2008, Chapter 7.

11 Samimi, A., A. (Kouros) Mohammadian, and K. Kawamura. A behavioral freight movement microsimulation model: method and data. Transportation Letters: The International Journal of Transportation Research (2010) 2: (53-62).

For the mesoscale model an adaptation of the ADA formulation is used. In this approach, total transport and logistics cost is evaluated as described by the ADA framework. Logistics options in this formulation include the following:

- Frequency or shipment size;
- Loading unit choice (e.g., container or trailer);
- Use of intermediate handling facilities such as distribution centers and intermodal yards (note: the use of intermediate facilities is modeled in the ADA framework);
- Mode used for each leg of the trip;
- Inventory costs; and
- Miscellaneous other costs such as ordering, damage, deterioration, and stockout costs.

The logistics-cost equation that is proposed in the ADA framework is:

Total annual logistics costs (G) (by commodity type) =

order costs (O)

+ transport and intermediate handling costs (T)

+ deterioration and damage costs (D)

+ capital costs of goods during transit (pipeline costs Y)

+ inventory costs (I)

+ capital costs of inventory (K)

+ stockout costs (Z)

If the error is assumed to be extreme value-distributed, then the model becomes a multinomial logit model where the choices are essentially individual paths that are associated with a unique "bundle" of transport and logistics attributes such as mode, shipment size, and number of intermediate handling facilities used. The full-cost function can be parameterized as shown in Equation 1.1 with the notation described in Table 1.1.

Equation 1.1 Parameterized-Cost Function for Multinomial Logit Model

$$G_{mnql} = \beta_{0ql} + \beta_1 * (Q / q) + T_{mnql} + \beta_2 * j * v * Q + \beta_3 * t_{mnl} * j * v * Q / 365 + (\beta_4 + \beta_5 * v)(q / 2) + a * Sqrt(LT * \sigma_Q^2 + Q^2 * \sigma_{LT}^2)$$

While the development of the prototype mesoscale model does not include a model estimation effort (nor are data available for this effort), the model coefficients should eventually be calibrated based on observed choice data. For example, survey data will be needed to calibrate the parameters of this model. It also is worth noting that because this cost function has non-linear entries, Ben-Akiva and de Jong used the Box-Complex iterative method for calibrating the parameter values.

Table 1.1 Description of Variable and Parameter Notation

Variable or Parameter	Description or Interpretation (of Parameters)	Source
G_{mnql}	Logistics cost between shipper m and receiver n with shipment size q and logistics chain l	Calculated in mesoscale model
Q	Annual flow in tons	Macroscale model or FAF
q	Shipment size in tons	Variable
β_{0ql}	Alternative-specific constant	Parameter to be estimated
β_1	Constant unit per order	Parameter to be estimated
T	Transport and intermediate handling costs	Lookup table for prototype; network skims for full scale model, probably informed by survey data
β_2	Discount rate	Parameter to be estimated
j	Fraction of shipment that is lost or damaged	Survey data or assumed value
v	Value of goods (per ton)	Calculate using FAF data
β_3	Discount rate of goods in transit	Parameter to be estimated
t	Average transport time (days)	Lookup table (or skims), possibly informed by survey data
β_4	Storage costs per unit per year	Parameter to be estimated
β_5	Discount rate of goods in storage	Parameter to be estimated
a	Constant used to set the safety stock in a way that generates a fixed probability of not running out of stock	Survey data or assumed value
LT	Expected lead time (time between ordering and replenishment)	Lookup table (or skims), possibly informed by survey data; total travel time plus time for order to be filled
σ_Q	Standard deviation in annual flow (i.e., anticipated variability in demand)	Survey data, assumed value, macroscale model, or other source
σ_{LT}	Standard deviation of lead time	Lookup table (or skims), possibly informed by survey data

The cost function described here presents a fairly comprehensive accounting of total logistics costs. All elements of this formulation are maintained in the prototype. However, during the full implementation, the model may only use the more important elements, such as transport and intermediate handling costs and shipment size, and simplify the less critical ones. In contrast, inventory costs (both pipeline and storage), loss and damage (L&D) costs, and safety stock costs are included in the formulation but might not be feasible to collect in surveys. If this is the case, then these elements may be included in a simplified form.

In addition to the ADA framework elements, market segmentation is used in the mesoscale prototype to better explain the effects of business or goods characteristics on path selection. For example, agents who trade bulk commodities are assigned parameters that effectively increase the appeal of using rail or water modes.

■ 1.4 Summary

The prototype model is designed as an analytical tool for understanding freight-related mode and logistics choices. The key elements of the mesoscale model framework include the following:

- A focus on commodity moves;
- The shipper-receiver or supplier-buyer pair as the key decision-making agent;
- Supply chain formation using high-level decision rules;
- A zone system comprised of townships in the seven-county area and larger areas outside of the seven-county area;
- A regional freight network, including transport and logistics nodes;
- National truck and rail network links;
- Regional water network links;
- Modeling freight moves through logistics nodes such as airports, water ports, intermodal yards, and distribution centers; and
- Evaluation of paths using the total transport and logistics costs for each path.

2.0 Data

This section describes the data inputs that were used for the prototype model and their sources. Future data requirements also are discussed.

2.1 Prototype Inputs

Table 2.1 lists a summary of data inputs that are required for the prototype mesoscale model. This table lists each data input and describes its source, the module(s) where it is applied, and a general description of the data. These inputs are described further in this section.

Table 2.1 Data Inputs

Type of Input	Input	Source	Module	Description
Zones	FAF3 Zone System	FAF3	Firm generation, Supplier selection, Flow apportionment	Large regions such as Combined Statistical Areas (CSAs) or states
	CBPZONE Zone System	Created by project team	Firm generation, Supplier selection, Flow apportionment	Counties (within the CMAP region) and FAF3 zones (outside of the CMAP region)
	FAFchi Zone System	Created by project team	Flow apportionment	Groups of counties (within the CMAP region) and FAF3 zones (outside of the CMAP region)
	Mesozone Zone System	Created by project team	Path selection	Townships or counties (within the CMAP region) and FAF3 zones (outside of the CMAP region)
Network Elements	Network links	Created by project team	Path selection	Highway, rail, and water network links
	Transport and logistics nodes (TLN)	Created by project team	Path selection	Specific nodes within the CMAP region; representative nodes outside of the region
	Rail service	Created by project team	Path selection	Rail routes with carrier identified

Table 2.1 Data Inputs (continued)

Type of Input	Input	Source	Module	Description
Skims	Great Circle Distance (GCD)	Oak Ridge National Laboratory (ORNL) County-to-County Matrix	Supplier selection	Distance between all county-level O-D pairs in the U.S.
	Foreign Skims	Created by project team	Supplier selection	Distance between U.S. counties and foreign FAF zones
	Path Skims	Created by project team in EMME/3	Path selection	Distances and logistics nodes used
Correspondences	Input-Output Make and Use Tables	U.S. Bureau of Economic Analysis (2002)	Supplier selection, Flow apportionment	Values of commodities exchanged between industries
	Industry to Commodity Correspondence	Freight Activity Microsimulation Estimator (FAME)	Firm generation, Supplier selection, Flow apportionment	List of SCTG commodities produced by each NAICS6 industry
	NAICS6 Industry to Input-Output Industry Correspondence	U.S. Bureau of Economic Analysis (2002)	Supplier selection, Flow apportionment	Correspondences between detailed NAICS6 industry classes and aggregated NAICS Input-Output industry classes
Flow Data	FAF3 Flows	FAF3	Supplier selection, Flow apportionment	Commodity flows between FAF3 zones
Employment Data	County Business Pattern (CBP) Data	U.S. Census (2007)	Firm generation, Supplier selection, Flow apportionment	Employment by industry
	Agricultural Employment	Created by project team	Firm generation, Supplier selection, Flow apportionment	Employment in agricultural industries
	Foreign Employment	Created by project team	Firm generation, Supplier selection, Flow apportionment	Employment by industry in foreign FAF3 zones
	Subzone Employment	CMAP	Flow apportionment, Business location	Employment in CMAP region

2.1.1 Zones

The model stream uses four zone systems. The different systems are used for apportioning high-level commodity flows to individual shipper-receiver pairs and identifying the set of feasible transport paths for each shipper-receiver pair. The four systems are as follows:

1. The broadest zone system, which is comprised of FAF3 zones (Figure 2.1), is used for the commodity flow input data. These zones also are used as "mesozones" in the final mesozone zone system to represent zones outside of the CMAP region.

2. An intermediate zone system comprised of "CBPZONE" zones (Figure 2.2) is used during several model processes, including **Firm Generation** and **Supplier Selection**. These zones consist of counties in the CMAP region and FAF3 zones outside of the CMAP region. Two exceptions include: the FAF3 Milwaukee zone and the FAF3 Northwest Indiana zone, which are smaller than their original FAF3 counterparts due to removal of CMAP counties from these two areas).

3. The FAFchi zone system consists of zones that are FAF3 zones outside of the region and groups of CBPZONE zones within the region (Figure 2.3). This zone system is used during **Flow Apportionment**.

4. Finally, firms in the CMAP region are assigned to Mesozones (Figure 2.4), which are FAF3 zones outside of the region, county-sized zones on the fringes of the region and township-sized zones in the inner counties.

Figure 2.1 Freight Analysis Framework (FAF3) Zones

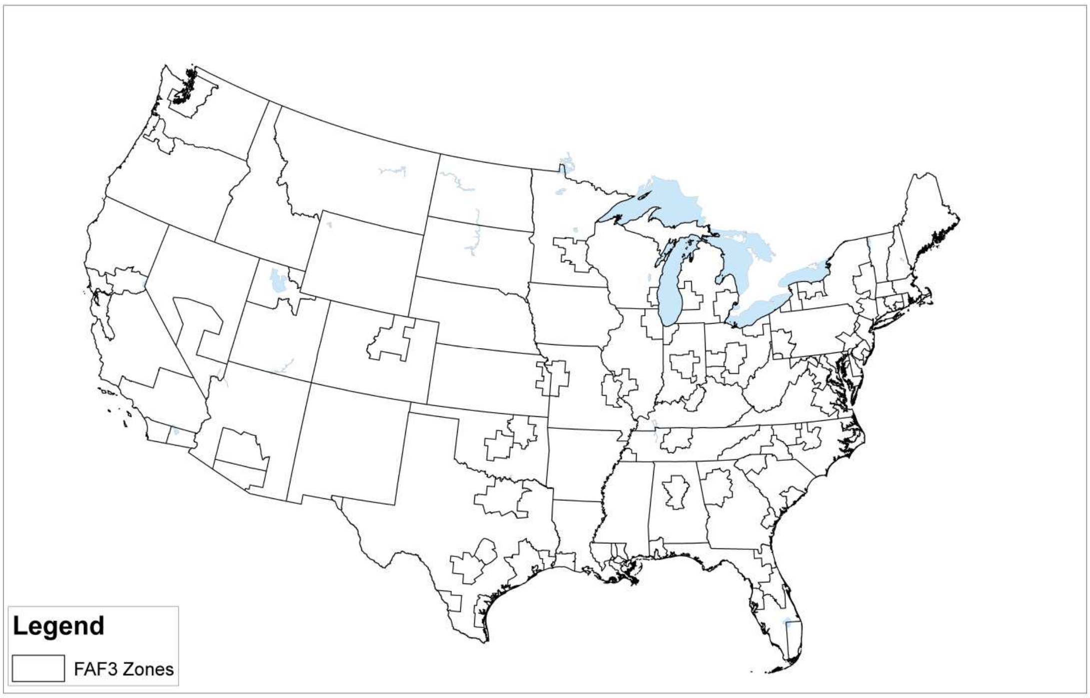


Figure 2.2 Intermediate Zone System (CBPZONE Zones)

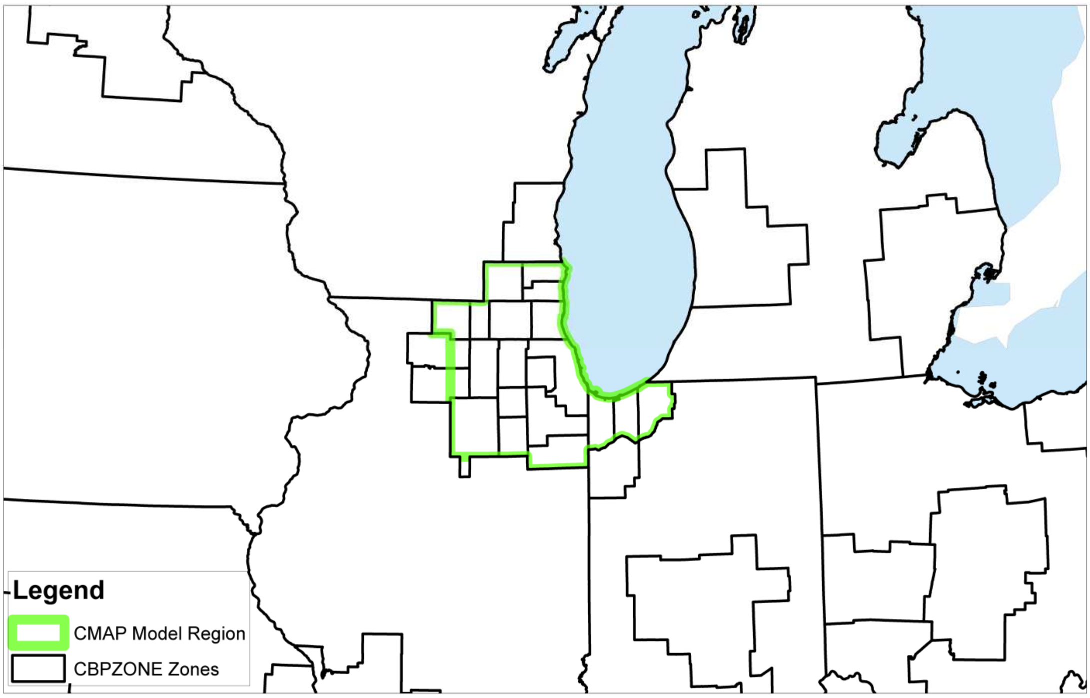


Figure 2.3 Intermediate Zone System (FAFchi Zones)

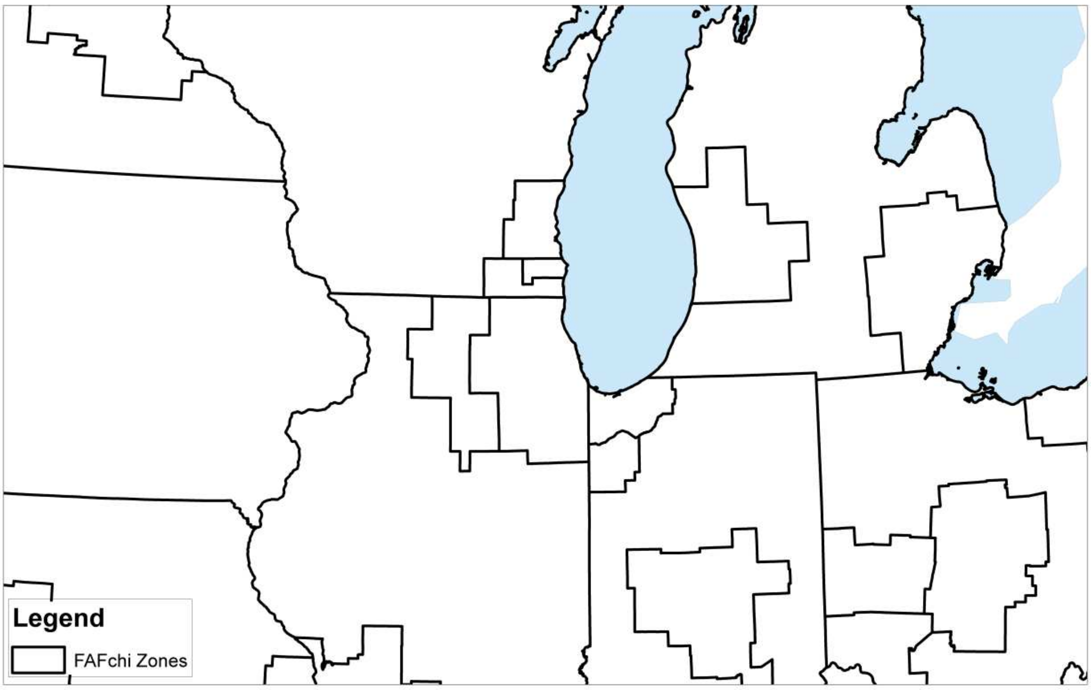
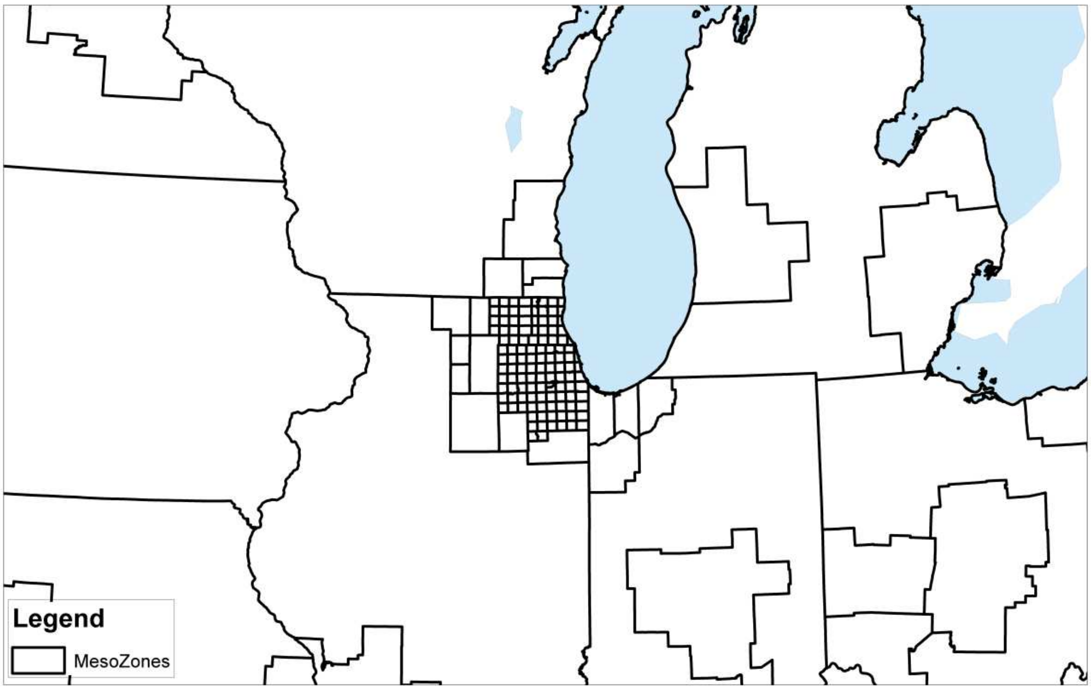


Figure 2.4 Mesoscale Zones



2.1.2 Network Elements

The prototype network is housed in EMME/3. Rail links, highway links, water links, and major logistics nodes are present both nationally and regionally. Figure 2.5 shows the entire network and Figure 2.6 shows a close-up of the regional network.

Figure 2.5 National "Stick" Network

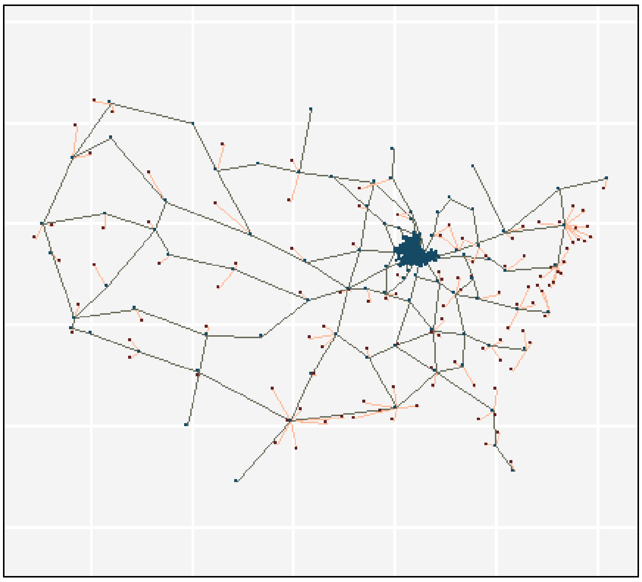
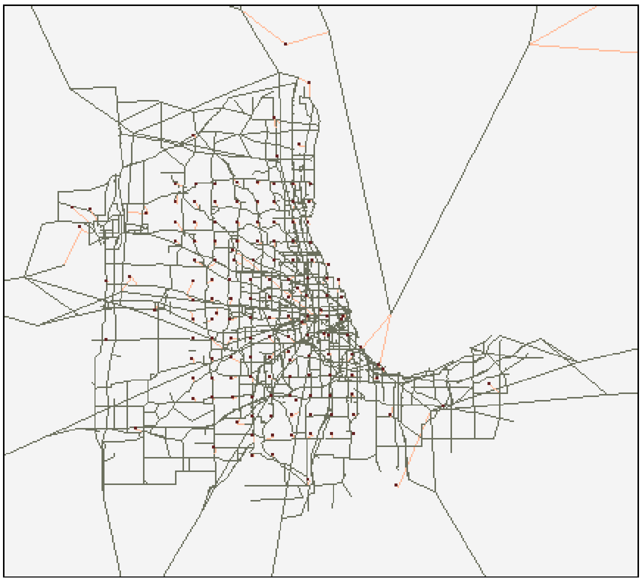


Figure 2.6 Regional Freight Network



2.1.2.1 Links

The network links include:

- Truck routes in the CMAP region;
- Links between the regional highway network and the national rail network;
- Links between external zone centroids and the national network;
- A generalized Class I rail and interstate highway network outside of the region; and
- A simplified network of water links in the Great Lakes region.

The rail network is based on the Oak Ridge National Laboratory (ORNL) "QC93L" network, which was last revised in 2009 and only includes rail links that currently are operational.[12] For the mesoscale prototype, a basic, functional, and coherent national rail network was developed based on the ORNL Class I railway network using travel demand modeling software tools. The network includes more detail for links that are radial to the Chicago area.

Because the national interstate network parallels the Class I network closely, the same national network was used to generate both highway and rail freight distance skims.

Most of the regional network links represent major truck routes in the CMAP region. However, the regional network also contains exemplary rail links that provide service between specific nodes (such as intermodal yards) and the national Class I service network. Figure 2.7 shows an example of these access links. The letters on each link represent the mode that is allowed to travel on the link. "T" stands for truck; the other modes are described in Table 2.2.

[12]Source: Oak Ridge National Laboratory website. Accessed on March 17, 2011 at: http://cta.ornl.gov/transnet/rrdescr.txt.

Figure 2.7 Close-Up of Rail Terminals and Linkages to the Rail Network (*Logistics Nodes 147 and 148 Are Shown)*

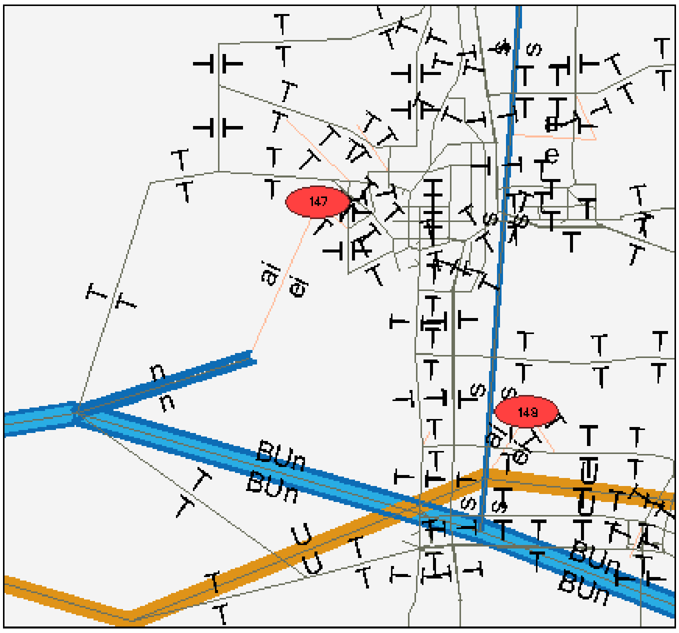
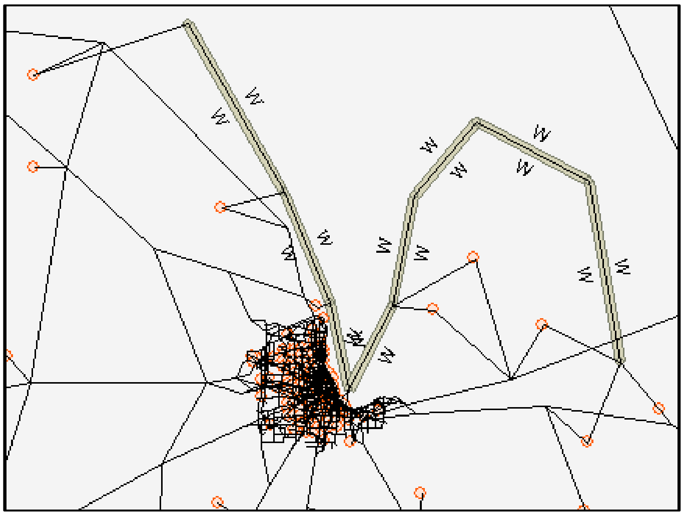


Table 2.2 Number of Rail Lines by Operator in Prototype Network

Mode Code (in EMME/3)	Code Description	Number of Lines in EMME/3
B	BNSF	8
n	Canadian National	8
p	Canadian Pacific	4
X	CSX-T	6
K	Kansas City Southern (KCS), Texas Mexican (TM), Gateway Western Railway (GWWR)	1
N	Norfolk Southern	7
U	Union Pacific	9
s	Short Line	2
a	Access mode	N/A
e	Egress mode	N/A

Figure 2.8 shows the rudimentary Great Lakes water network that was developed for the prototype. This network does not cover all possible Great Lakes water movements; instead, it functions as a placeholder for a more detailed network (to be developed in the full model implementation).

Figure 2.8 Water Network



Distance information from the national network is processed to generate Level of Service skims. The network can be further embellished to include other factors such as link-specific travel times.

2.1.2.2 Transport and Logistics Nodes

The network nodes include both zone centroids and logistics nodes. Logistics nodes play two critical roles in the prototype model. First, all of the logistics nodes are used to generate handling activities and their associated costs. Second, logistics nodes within the CMAP region in combination with the path generation methodology provide a way to evaluate and compare locational advantages among facilities. For example, this methodology can be used to compare preferences among shipper-receiver pairs for using different airports. This and other model applications are described further in Section 4.3

For the prototype, a sample of major logistics nodes within the region was developed. Each of these nodes corresponds to one facility. Table 2.3 lists the logistics nodes that are featured in the regional network and gives a brief description of each. Zone numbering ranges are also documented in this table. Figure 2.9 shows the rail terminals that are coded in the prototype network; Figure 2.10 shows the airports; Figure 2.11 shows the water ports; and Figure 2.12 shows the truck terminals.

Table 2.3 Logistics Nodes and Zones Featured in the Prototype Network

Type	Node	Description
Zone	1-132	Mesozones in CMAP region
Truck terminals	133	Kenosha-area terminal (e.g., JHT Holdings, Inc.)
	134	Rockford-area terminal (e.g., UPS Freight)
	135	DuPage/Kane-area terminal (e.g., Cassens Transport Co.)
	136	Will County terminal (e.g., Packard Transport, Inc.)
	137	Hammond/Gary area terminal (e.g., Traco Transportation)
	138	Terminal near I-294/I-290 junction (e.g., Transport Service Co.)
	139	Central Chicago terminals
Unused	140	
Airports	141	O'Hare
	142	Midway
	143	Gary
	144	Milwaukee
Water ports	145	Illinois International Port
	146	Indiana Harbor
Rail terminals	147	Rockford-area yards
	148	Global III – Rochelle
	149	Logistics Park – Elwood
	150	Central Chicago yards
Zone	151+	Mesozones outside of CMAP region (domestic)

Figure 2.9 Regional View of Rail Network, Routes, and Rail Terminals (Logistics Nodes 147-150)

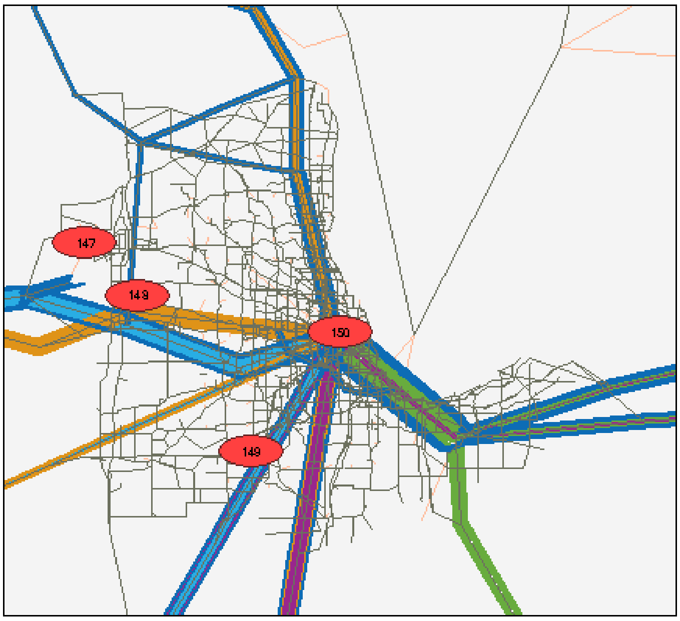


Figure 2.10 Regional Airports: Milwaukee, O'Hare, Midway, Gary (Logistics Nodes 141-144)

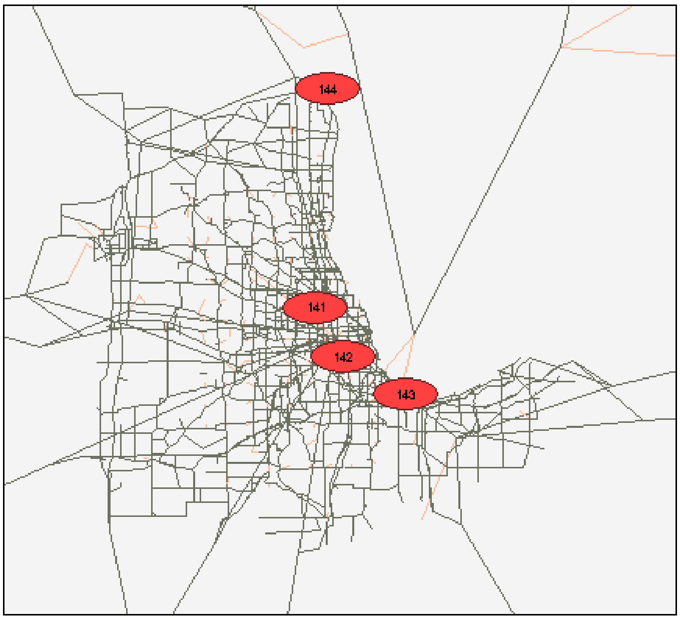


Figure 2.11 Water Ports (Logistics Nodes 145-146)

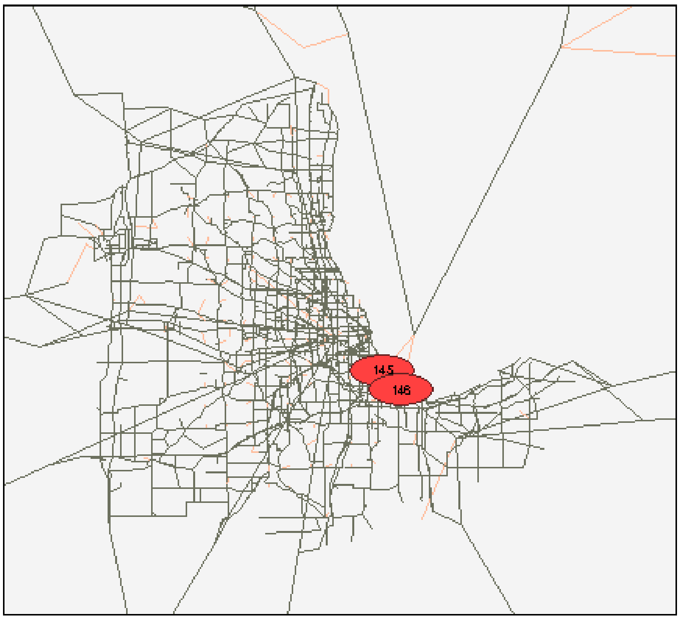
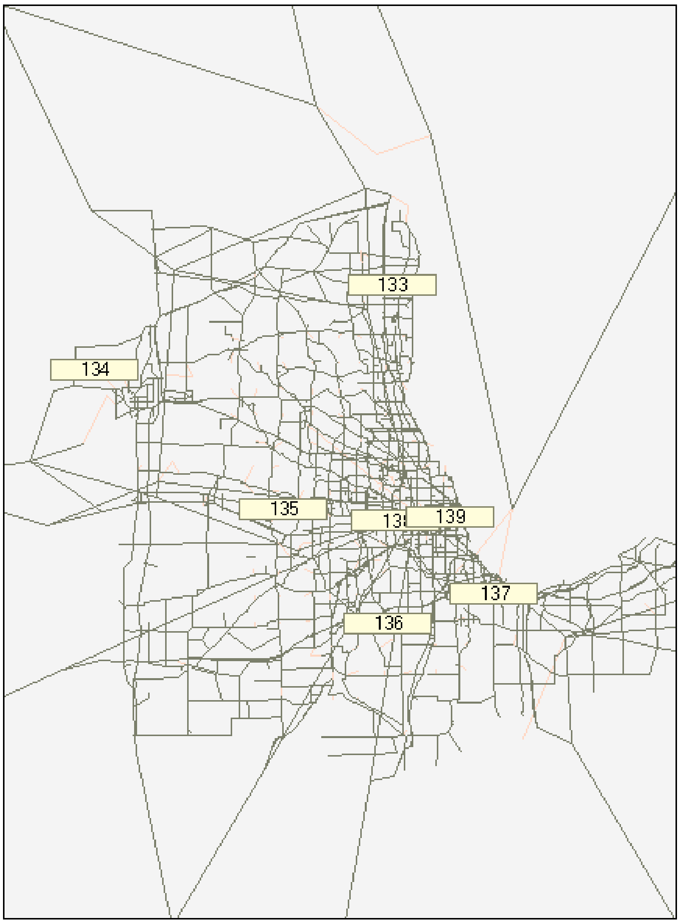


Figure 2.12 Truck Terminals (Logistics Nodes 133-139)



At the national level, the purpose of these nodes is to provide a convenient framework for generating "transfer penalties" that account for the cost of performing an intermediate handling move. Accounting for other details for nodes outside of the Chicago region is not necessary since, each category of facilities in an external zone is represented by a single logistics node irrespective of the actual number or locations of individual facilities. These logistics nodes are used to represent a variety of logistics handling activities during the model processing. This means that although only one node is used generically during skimming to represent all types of logistics handling outside of the region, the actual transport and logistics-cost calculation methodology is flexible enough to accommodate any number of handling facility types and costs.

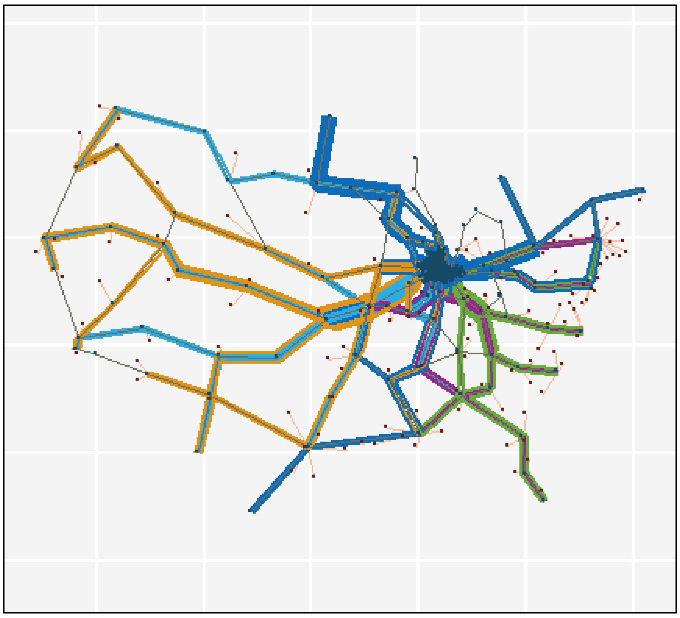
2.1.2.3 Rail Service

The Class I rail operators treat their operating information, including routes operated, as private business information. For this reason, actual Class I routes were not available for coding into the EMME/3 model. Instead, hypothetical Class I routes were developed for the mesoscale model based on information from the ORNL network (Figure 2.13).

Because they are not based on observed route data, they essentially represent service availability by operator between two zones but do not represent actual travel times or routes taken.

Each rail carrier was assigned a unique mode for route development and skimming using the EMME/3 network. Table 2.2 shows the mode code for each carrier and the number of rail routes that are coded into the prototype model.

Figure 2.13 National Rail Network and Routes



Two "short-line" rail services, which frequently facilitate movements between local businesses and the national Class 1 routes, also are included in the prototype network. These currently are not being used by the model but are included for consideration in the full implementation of the model. To include them in the fully implemented model, more specific details on routes, transfer points, and alignment should be determined. Then, these relatively detailed short-line routes could be modeled using EMME/3 to generate both short- and long-distance skims that involve a short-line move.

Alternatively, as was considered for the prototype, the long-haul skims can be adjusted to include information on short-line trips.

Figure 2.9 shows a regional overview of the rail routes and several rail-related logistics nodes. These nodes represent intermodal yards and break-bulk facilities.

2.1.3 Skims

Great circle distances for the CBPZONE zone system are used during the **Supplier Selection** stage. Domestic distances were obtained from ORNL county-to-county distance matrices. Distances between Chicago and foreign FAF3 zones were estimated using free on-line distance lookup tools.

Great circle distance also was used to compute level of service for paths between the CMAP area and foreign zones in the **Path Selection** phase. For international flows that use overseas water and a domestic ground mode, distance between various port cities (Los Angeles, Seattle, New York/New Jersey, and New Orleans) were used to compute skims for the domestic portion of the trip.

Domestic highway, rail, and water skims were generated using EMME/3 skimming procedures. Additional distances were added to generate distances to and from Canada and Mexico. Since highway and rail skims were obtained using a network that is mostly radial to Chicago, domestic air distance was considered to be approximately the same as highway distance. Table 2.4 shows how the available skims were analyzed to develop skims for foreign and non-contiguous U.S. areas. The "GCD (Great Circle Distance) to Chicago" field contains a value only if it is used to represent distance in this process. Otherwise, distance is generated during the EMME/3 skimming procedures.

Table 2.4 International and Non-Contiguous U.S. Skims

FAF3 Zone ID(s)	Region	Zone of Entry/Exit	Mesozone ID	Added Distance (miles) from Zone of Entry/Exit	GCD to Chicago	Allowed Modes	Notes
801	Canada	Minnesota	206	500	–	All	
802	Mexico	Houston	257	300	–	All	
803	Rest of Americas	New Orleans	195	N/A	3,000	Air or Water and Ground	Water and Ground path has a transloading option
804	Europe	New York	217	N/A	4,000	Air or Water and Ground	Water and Ground path has a transloading option
805	Africa	New York	217	N/A	7,000	Air or Water and Ground	Water and Ground path has a transloading option

Table 2.4 International and Non-Contiguous U.S. Skims (continued)

FAF3 Zone ID(s)	Region	Zone of Entry/Exit	Mesozone ID	Added Distance (miles) from Zone of Entry/Exit	GCD to Chicago	Allowed Modes	Notes
806	South, Central and Western Asia	Seattle	268	N/A	8,000	Air or Water and Ground	Water and Ground path has a transloading option
807	East Asia	Los Angeles	159	N/A	7,000	Air or Water and Ground	Water and Ground path has a transloading option
808	Southeast Asia and Oceania	Los Angeles	159	N/A	9,000	Air or Water and Ground	Water and Ground path has a transloading option
151, 159	Hawaii	N/A	N/A	N/A	4,250	Air	
20	Alaska	North Dakota	230	2,700	–	Air or Ground	

2.1.4 Correspondences

2.1.4.1 Input-Output Make and Use Table

The Bureau of Economic Analysis' Input-Output Make and Use table is used to create supply chains in the **Supplier Selection** component. For each production industry, the table reports the value of goods consumed by each buyer industry. The model uses this information to identify for each buyer industry the most important commodities that are consumed and their associated supplier industries.

The reported values of goods exchanged between producers and consumers also are used to apportion FAF3 flows by commodity type. Two types of apportioning take place. First, flows from the FAF3 zone system are divided between the portion of the FAF3 zone that is in the CMAP region and the portion that is outside of the region. Second, the flows that result from the geographic redistribution are apportioned among the supply chain pairs. These apportionments use information from the Make and Use table to determine the total volume of a commodity that is produced or consumed by a particular industry compared to other industries that produce or use the commodity.

2.1.4.2 Other Correspondences

The prototype model employs several correspondence files to relate certain data items to one another. For example, correspondences between the different zone systems were established. Other correspondence files report the commodity (or commodities) produced and/or consumed by certain industries. Finally, a key correspondence file relates the detailed NAICS6 industries to the broader industry categories used in the Make and Use tables.

2.1.5 Macroscale Freight Flows

Macroscale freight flow data contain information on commodity movements between the CMAP region and other domestic and foreign regions. The prototype model relies on FAF3 data, which reports commodity flows in units of tons per year and value (in USD) per year using the two-digit SCTG system.

The freight flow data are used in two ways. First, the origin and destination pairs reported in the FAF3 data are used to identify candidate suppliers for every buyer during **Supplier Selection**. Second, the flow data are apportioned to individual supplier-buyer pairs.

2.1.6 Employment Data

County-level employment data for the entire U.S. are used to generate firms. For each county, this dataset contains the number of firms in each category defined by industry (six-digit NAICS) and number of employees. The prototype relies on County Business Patterns (CBP) data. However, the CBP data contain limited or no agricultural, construction, or foreign employment. As a result, the project team developed data in these areas to be used as placeholders until sufficient data can be obtained. This effort is described below.

Employment data also are used in conjunction with information from the Make and Use table to apportion FAF3 flows. The Make and Use table provides the value of goods that are traded between different industries. The resulting values are effectively weighted by employment for apportioning the flows.

In addition to county-level data, the model used employment data for subzones in the CMAP region. These data are used to assign each firm in the CMAP area to a specific mesozone using a rudimentary business location model. Subzone employment data were obtained from CMAP.

2.1.6.1 Development of Agricultural Employment Data

Agricultural employment data were developed using FAF3 production totals for agricultural commodities and supplementary data from sources such as the U.S. Department of Agriculture. Appendix A documents this effort. Total employment for each NAICS agricultural industry was then allotted to the eight business size categories. For the prototype, it was assumed that the distribution of size among agricultural business is the same as the distribution of size for all businesses.

2.1.6.2 Development of Construction Employment Data

Construction employment for the Chicago area was generated using CMAP subzone employment for NAICS industry 23. Total NAICS 23 employment by county was allotted to the seven construction industries that appear in the Make and Use tables based on the average payroll by industry, which also is based on data from the Make and Use table. Table 2.5 lists the six-digit NAICS industry IDs and the derived percentage that was used to allocate total construction employment to each subindustry.

Table 2.5 Construction Industries in the Make and Use Table

NAICS	Percentage of NAICS 23
230101	15%
230102	3%
230103	31%
230201	25%
230202	11%
230301	13%
230302	2%

In addition, construction employment was generated for non-CMAP zones. The purpose was to ensure that producers of construction-related commodities such as gravel that are located in the CMAP region would be able to locate buyers outside of the region. This is important for allocation of FAF3 flows.

A simple approach was adopted to generate construction employment outside of the CMAP area. For the prototype model, it is assumed that every non-CMAP zone has the same distribution of construction industries by size. For each external zone, two businesses are generated for each six-digit NAICS industry (Table 2.5) and each employment size category.

Table 2.6 shows the resulting number of construction businesses by size and geographic area. The number of firms by size is shown for each of the counties in the CMAP region. For firms that are located outside of the region, the total number of firms by size is shown. Because construction firms outside of the region were generated uniformly by size category and region for the prototype, the total number of external firms is the same for each size category.

Table 2.6 Number of Construction Businesses by Size Generated for the Mesoscale Model

County (CMAP Region)	Number of Employees 1 to 19	20 to 99	100 to 249	250 to 499	500 to 999	1,000 to 2,499	2,500 to 4,999	Over 5,000
17007	27	4	–	–	–	–	–	–
17031	495	84	29	14	9	6	5	2
17037	20	2	–	–	–	–	–	–
17043	271	46	17	11	5	2	2	–
17063	22	3	–	–	–	–	–	–
17089	146	25	7	6	2	–	–	–
17091	33	6	1	–	–	–	–	–
17093	25	3	–	–	–	–	–	–
17097	196	32	10	8	3	–	–	–
17099	17	2	–	–	–	–	–	–
17103	–	–	–	–	–	–	–	–
17111	117	20	6	3	1	–	–	–
17141	5	–	–	–	–	–	–	–
17197	183	30	10	8	3	–	–	–
17201	104	17	6	3	1	–	–	–
18089	207	35	12	9	4	–	–	–
18091	113	19	7	2	–	–	–	–
18127	107	18	7	2	–	–	–	–
55059	84	14	6	2	–	–	–	–
55101	100	15	6	2	–	–	–	–
55127	74	12	4	1	–	–	–	–
Total Firms, CMAP	2,346	387	128	71	28	8	7	2
Total Firms, Elsewhere	1,708	1,708	1,708	1,708	1,708	1,708	1,708	1,708

2.1.6.3 Development of Foreign Employment Data

The primary objective of including foreign firms in the mesoscale model is to ensure that international flows between the Chicago area and foreign countries will be properly processed by the mesoscale model. For the prototype, this was accomplished by generating agents for all types of industries and commodities in the foreign FAF3 zones. Foreign firms were generated by assuming that all foreign regions have firms in each NAICS6 industry.

2.2 Data Needs

2.2.1 Network Enhancements

A simplified network was developed for the prototype. Other elements that can be implemented with more detail include:

- National Highway System (NHS) intermodal connectors;
- Speed limit;
- Functional class;
- Truck counts by class (to use in validation);
- Information on restrictive overhead clearances or bridges with weight restrictions;
- A more detailed rail network;
- Actual rail routes;
- Rail Level-of-Service (LOS) indicators;
- More detailed water links/expanded water network;
- Links that connect rail lines of different ownership; and
- Additional logistics nodes in the Chicago region, including:
  - Intermodal yards;
  - Container yards[13] (these overlap somewhat with intermodal yards);
  - Terminal facilities, including additional air and water ports. It may be useful to code in a sample of these facilities or a single node to represent a cluster of these facilities at the township level;
  - Major distribution centers and warehouses;[14] and
  - Businesses with a regionally significant freight presence such as large retail malls and factories.

---

[13]CS and CMAP distinguished between container yards and intermodal yards for the CMAP Freight Plan Recommendations study. The same categorization is recommended for the mesoscale model.

[14]CMAP already has a shapefile of regional distribution center (DC) locations. Regionally important warehouses function essentially as DCs, therefore it is not necessary to collect any additional data on warehouses. Smaller warehouses, if needed for modeling, will be identified using employment data.

2.2.2 Employment Data

For the prototype model, placeholder data representing construction, agricultural, and foreign businesses were developed. The methodology used for developing construction and foreign employment was simplistic. The methodology used to develop agricultural business data was more rigorous but has not been thoroughly tested. These datasets can be improved upon for the fully implemented mesoscale model.

Furthermore, future-year employment data that resemble the base-year CBP data would be needed for future-year applications of the model. For example, this dataset can be created by applying a flat growth rate to the entire CBP dataset.

2.2.3 Parameters and Assumptions

Parameters and assumptions are used in three critical areas of the model. First, they are used to generate level of service data. Second, they are used in the **Supplier Selection** component. Third, they are used in formulation of path costs. All assumptions and parameters that are used in the prototype should be revisited and/or calibrated in the full implementation of the model.

2.2.3.1 Level of Service Parameters

Table 2.7 shows the Level of Service parameters that are being used in the prototype mesoscale model including handling times and delays at various logistics facilities. Total travel time is calculated based on the assumed access, egress and line-haul distances, and speeds. Total path cost is calculated as the sum of all transportation, handling, and other costs that are incurred.

Table 2.7 Level of Service Parameters

Parameter	Description	Value
BulkHandFee	Handling charge for bulk goods ($ per ton)	1.00
WDCHandFee	Warehouse/DC handling charge ($ per ton)	15.00
IMXHandFee	Intermodal lift charge ($ per ton)	15.00
TloadHandFee	Transload charge ($ per ton; at international ports only)	10.00
AirHandFee	Air cargo handling charge ($ per ton)	20.00
WaterRate	Line-haul charge, water ($ per ton-mile)	0.005
CarloadRate	Line-haul charge, carload ($ per ton-mile)	0.03

Table 2.7 Levels of Service Parameters (continued)

Parameter	Description	Value
IMXRate	Line-haul charge, intermodal ($ per ton-mile)	0.04
AirRate	Line-haul charge, air ($ per ton-mile)	3.75
LTL53rate	Line-haul charge, 53 feet LTL ($ per ton-mile)	0.08
FTL53rate	Line-haul charge, 53 feet FTL ($ per ton-mile)	0.08
LTL40rate	Line-haul charge, 40 feet LTL ($ per ton-mile)	0.10
FTL40rate	Line-haul charge, 40 feet FTL ($ per ton-mile)	0.10
WaterMPH	Water speed (mph)	5.00
RailMPH	Rail speed (mph)	30.00
LHTruckMPH	Line-haul truck speed (mph)	60.00
DrayTruckMPH	Drayage truck speed (mph)	45.00
AirMPH	Air speed (mph)	500.00
ExpressSurcharge	Surcharge for direct/express transport (factor)	1.50
BulkTime	Handling time at bulk handling facilities (hours)	72.00
WDCTime	Handling time at warehouse/DCs (hours)	24.00
IMXTime	Handling time at intermodal yards (hours)	24.00
TloadTime	Handling time at transload facilities (hours)	12.00
AirTime	Handling time at air terminals (hours)	1.00

Ideally, the Level of Service parameters eventually will be calibrated using observed data. While these assumptions are reasonable for the prototype, certain analysis situations may require a greater degree of accuracy.

This type of data already is available for at least some paths and modes (e.g., see Leachman's report[15] with average values between major U.S. ports and Chicago). However, more research needs to be conducted to assess the extent to which this type of data is available for what regions and for what level of detail.

[15] Leachman, R., T. Prince, T. Brown, and G. Fetty. Final Report: Port and Modal Elasticity Study. Southern California Association of Governments. September. 8, 2005. Accessed on-line on January 18, 2011 at: http://www.ieor.berkeley.edu/People/Faculty/leachman-pubs/PortModal.pdf.

2.2.3.2 Supplier Selection Parameters

The **Supplier Selection** model parameters that are used in the prototype are based approximately on the fuzzy logic parameters that are used in the **Supplier Selection** module of FAME. The FAME model assigns businesses to size categories (small, medium, large) based on the number of employees and it evaluates the Great Circle Distances between each consumer firm and potential supplier firms using categories (over 1,509 miles, 596-1,509 miles, and 595 or fewer miles).

Furthermore, in FAME's fuzzy logic formulation, probabilities are described in general terms: low, average, or high. For example, according to FAME, small consumer firms have a low probability of selecting a small producer firm, a low probability of selecting a medium sized producer firm, and an average probability of selecting a large firm to trade with. For the mesoscale prototype, low, medium, and high probabilities were translated into coefficients of 0.2, 0.4, and 0.6.

Similarly, distance parameters from FAME were adapted for use in the mesoscale model. As distance increases, the parameters become more negative to indicate decreased interaction between businesses. Distance categories were also adapted from FAME but were refined for the mesoscale prototype model to account for the refined zone system. The FAME model uses FAF3 zones for the entire modeling process whereas the mesoscale **Supplier Selection** module uses the CBPZONE system, which contains more detail than FAF3 zones within in the CMAP region.

Table 2.8 shows the parameters that are used in the prototype **Supplier Selection** model.

Table 2.8 Supplier Selection Parameters

Consumer Business Size (Number of Employees)	Coefficient							
	Producer Business Size (Number of Employees)			Great Circle Distance between Consumer and Producer (Miles)				
	1 to 99	100 to 499	500+	Over 1,509	596 to 1,509	150 to 595	1 to 149	0 (Intracounty)
1 to 99	0.2	0.2	0.4	-0.4	-0.3	-0.2	0	0.1
100 to 499	0.2	0.6	0.6	-0.2	-0.1	-0.05	0	0.1
500+	0.4	0.6	0.6	-0.1	-0.05	0	0	0.1

In addition to the coefficients shown in Table 2.8, the **Supplier Selection** formulation uses a random constant that is unique to each potential supplier. This ensures that a variety of suppliers will be selected (rather than the "best" one for each business size and distance combination).

Calibration of the **Supplier Selection** parameters will require collection and analysis of survey data. The complexity of the survey effort depends somewhat on the actual complexity of the decision-making process that is used by buyers when selecting a supplier.

2.2.3.3 *Path Cost Parameters*

Table 2.9 shows the path cost parameters that are assumed for the prototype. These parameters are used to calculate the annual transport and logistics path costs as described in Section 1. A total of 54 types of paths are evaluated in the prototype model.

Table 2.9 Path Cost Parameters

Parameter	Description	Value
B_1	Constant unit per order	50.00
B_4	Coefficient for discount rate	5,000.00
j	Fraction of shipment that is lost or damaged	0.01
a	Safety stock constant	0.50
LT_OrderTime	Expected lead time (time between ordering and replenishment)	10.00
sdQ	Standard deviation in annual flow	1.00
sdLT	Standard deviation in lead time	1.00
LowDiscRate	Low-discount rate	0.01
MedDiscRate	Medium-discount rate	0.05
HighDiscRate	High-discount rate	0.25
CAP1FTL	Truckload capacity (tons)	30.00
CAP1Carload	Carload capacity (tons)	32.00
CAP1Airplane	Air cargo hold capacity (tons)	1.00

Calibration of the path cost parameters will require collection and analysis of survey data. The effort that is expended in this exercise will directly impact the complexity of the path cost formulation that is adopted for the fully implemented mesoscale model. Table 1.1 in Section 1 contains the full list of cost components. As discussed before, some of these elements may be sacrificed in order to gain more detail on the "bigger picture" items such as mode choice and shipment size.

Potential survey options include revealed-preference or stated-preference surveys of shippers/receivers (or their intermediary, the carrier), GPS surveys where the paths of individual shipments are tracked, or some combination thereof.

3.0 Model Flow

The mesoscale model transforms macroscale high-level commodity flow data into commodity trips at a level of detail that is suitable for regional travel demand modeling. Four key processes are used in the model stream:

1. ***Firm Generation***: the synthesis of firms in the U.S. and abroad;
2. ***Supplier Selection***: the formation of supply chains between individual firms;
3. ***Flow Apportionment***: the apportionment of high-level commodity flow data to individual shipper-receiver pairs; and
4. ***Path Selection***: the selection of transport and logistics paths (including shipment frequency) that are used to transport shipments between shippers and receivers.

Previous sections described the inputs to each process; the objectives of each process; and how the individual components fit together from a conceptual standpoint. This section discusses how these processes have been implemented in a standard computing environment.

■ 3.1 Overview

The mesoscale model transforms high-level commodity flow data into an agent-based dataset that lists the origin, destination, and intermediate stops of shipments between each shipper-receiver pair. Information on mode and vehicle type for each leg of the trip is included. Stops at major logistics nodes are represented. The output data can be transformed into average weekday truck vehicle trips by commodity to be used in highway assignment.

This section provides a step-by-step description of how the model framework has been adapted to a computing environment. Pre- and post-processing steps also are described. Two computing environments are used for the model. The freight network and its skimming procedures are contained in EMME/3 and the remainder of the model uses SAS. The SAS/Access module is required to read in various *.dbf files and Access databases.

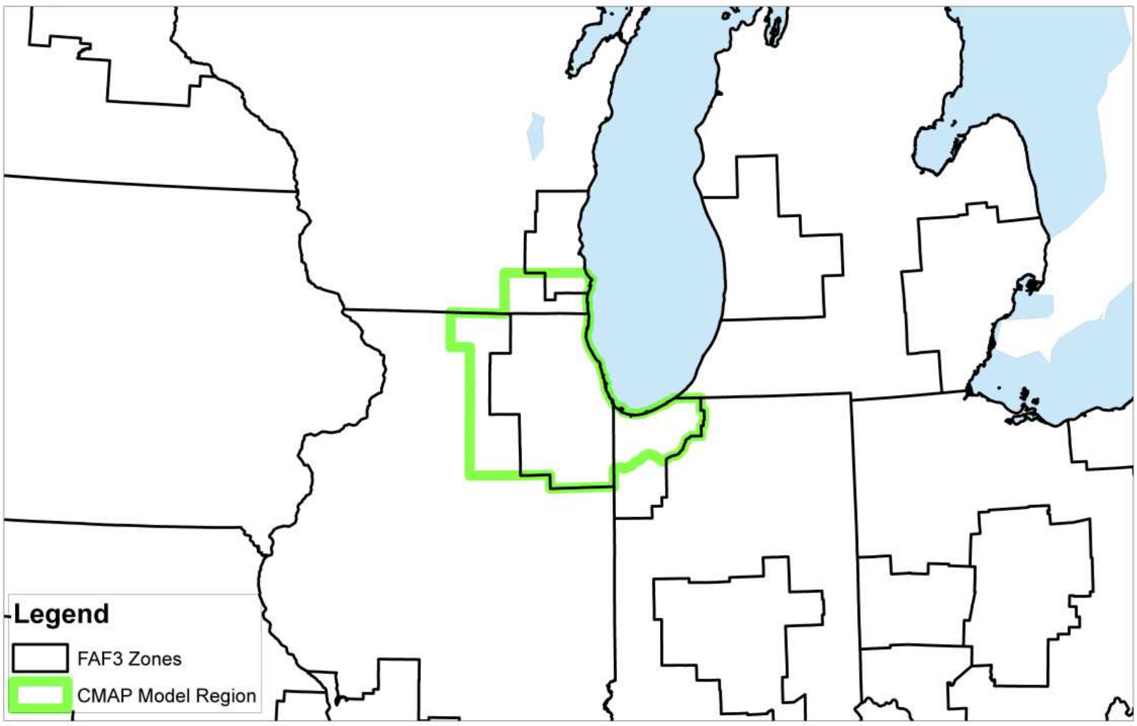
3.2 Preprocessing: Quantifying the Flow of Goods Moving Into, Out Of, and Through the Region

Macrolevel quantities must be provided by either the FAF base-year and forecast-year data or by the proposed macroscale model. For flows with one or both trip ends in the Chicago region, the FAF3 data provides to the prototype the volumes by commodity between production and consumption regions (Figure 1.1). The macroscale model also will be required to provide the mesoscale model with information on "through" trips. Additionally, the macroscale model probably will need to evaluate the costs of transport and logistics paths that go through the CMAP region versus other regions.

The ideal macroscale flow dataset will include a geographic region(s) that nests neatly with the CMAP region. However, the FAF3 zone system does not nest with the CMAP region. Instead, the CMAP region overlaps five different FAF3 zones. Figure 3.1 shows that:

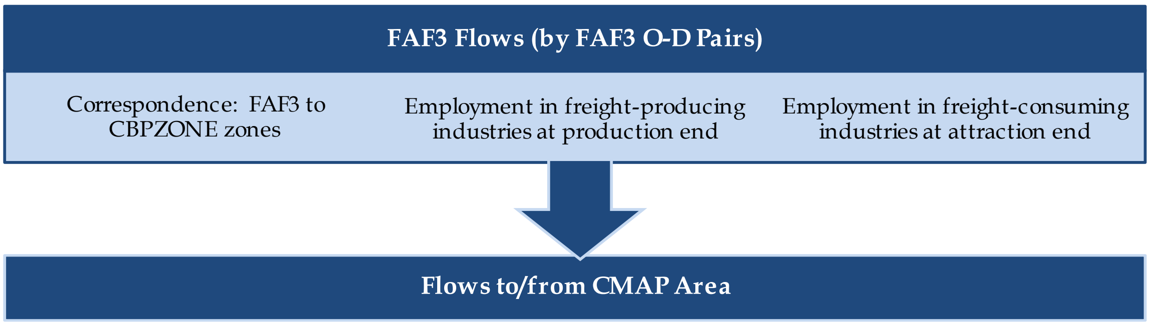
- The entirety of FAF3 region 171, the Illinois portion of the Chicago Community Statistical Area (CSA), corresponds to 10 CMAP counties; and
- The remaining portions of the CMAP form parts of four larger FAF3 zones.

Figure 3.1 CMAP Model Area and FAF3 Zones



The prototype mesoscale model refines the existing FAF3 origin and destination regions into a geography that is more appropriate for the CMAP region. For the four areas of overlap, the regional FAF3 flows are divided between CMAP and non-CMAP areas using employment data. Employment in industries that produce or consume commodities is used to apportion the flows. Industries are associated with commodities based on the NAICS-SCTG correspondence input table and the Make and Use table. Figure 3.2 presents a flowchart of the FAF3 flow customization process.

Figure 3.2 FAF3 Flow Customization



Finally, additional adjustments to the FAF3 data include:

- Imputing missing tons or values using average values from non-missing observations;
- Removing pipeline flows; and
- Removing flows that are reported to have a domestic origin or destination in the CMAP region but that actually originate and terminate in foreign countries.

The resulting commodity flows are used as inputs for the **Supplier Selection** and **Flow Apportionment** steps later in the model stream.

3.3 Firm Generation

3.3.1 Synthesis of Individual Firms

Individual firms are generated primarily using data on businesses that are maintained by the Census. The County Business Pattern (CBP) dataset contains the number of firms by size, industry, and county for every county in the U.S. Agricultural, construction, and foreign businesses are not included or are included in a limited fashion. Employment for these categories were developed for the prototype. Firms then are enumerated using these combined inputs.

Next, firms are prepared for use in the **Supplier Selection** stage. Firms are identified as Producers, Consumers or both. A correspondence table that was originally developed for FAME is used to establish the type of commodity that is produced by each firm (where applicable).

A unique ID is created for each firm and its firm-type is established. Producer firms are prepared for input as "firm-types" for the **Supplier Selection** stage. Firm-types are groups of businesses of similar size, industry class, and location. This substantially increases processing efficiency with no impact on the behavioral underpinnings of the model. The firm-type for each supply chain is redefined as a unique firm later in the model process. Consumer firms are maintained as uniquely identifiable firms throughout the selection process.

Producers (suppliers) and consumers (buyers) then are sampled from across the nation. The final sample that is input to the **Supplier Selection** module includes:

- All firms in the Chicago region;
- All of the largest firms elsewhere;
- At least one supplier firm for every possible firm-type combination;
- At least one buyer firm for every possible industry-zone combination; and
- A random sample of other buyer firms.

3.3.2 Modeling Firm Location

Finally, because **Path Selection** is performed using the mesoscale zone system, it is necessary to assign each firm to a specific mesozone. The finest level of geography that is available in the employment inputs is the county level. The **Firm Generation** process results in a list of individual firms that are identified using the CBPZONE system, which is comprised of counties (for firms that are located in the CMAP region) or FAF3 zones (for firms that are outside of the CMAP region). For external firms, the mesozone is the same as the CBPZONE, although renumbering takes place. For firms in the collar counties of the CMAP region, the mesozone is the same as the county or CBPZONE. Other firms in the CMAP region must be assigned to a specific mesozone for the **Path Selection** module.

A rudimentary location model is used to assign firms to mesozones in the seven-county CMAP area. The primary objective of this process is to assign the largest businesses to the mesozones that have the most employment in the relevant industries. This process involves the following steps:

1. Subzone employment data by industry are aggregated to a) the mesozone level and b) the county level;

2. Within each county, each mesozone is assigned to one of ten percentile ranking categories. Thresholds of 10 percent are used to determine the categories. Table 3.1 provides an illustration of this process for mesozones with NAICS 31 to 33 in Lake County. As a result of this process, mesozones with the highest employment are classified as Rank 10, zones with the lowest employment Rank 1, and so on;

3. Firms that were generated in the mesoscale model are assigned to a mesozone in the county in which the firm is located. Firms with more than 5,000 employees are randomly assigned to mesozones that have very-high employment in the firm's industry class (i.e., mesozones with Rank 9 or 10; Table 3.1). Firms with between 2,000 and 5,000 employees are randomly assigned to mesozones that have relatively high employment in the firm's industry (Rank 7 or higher); firms with 500 to 2,000 employees are assigned to zones with Rank 5 or higher; firms with 100 to 500 employees are assigned to zones with Rank 4 or higher; firms with 20 to 100 employees are assigned to zones with Rank 2 or higher; and firms with 20 employees or fewer to any zone with employment in its industry.

This procedure is intended to prevent the assignment of a relatively large firm to a zone with very little employment in the firm's industry class. It also prevents assignment of firms to zones that do not have employment in the correct industry. This procedure may be further improved in the full model implementation. For example, for the fully implemented model, it may be worthwhile to assign firms to the individually geocoded business locations that are used to build the subzone employment dataset.

Table 3.1 Percentile Ranking for Mesozones in Lake County, Illinois: *NAICS 31 to 33*

County	Mesozone	Employment	Percentage of Employment in County	Cumulative Percent	Percentile Rank
17097	46	27	0.0%	0.0%	1
17097	48	171	0.3%	0.3%	1
17097	49	260	0.4%	0.8%	1
17097	50	302	0.5%	1.3%	1
17097	58	552	0.9%	2.2%	1
17097	47	584	1.0%	3.2%	1
17097	57	901	1.5%	4.8%	1
17097	55	902	1.5%	6.3%	1
17097	51	1,124	1.9%	8.2%	1
17097	54	1,165	2.0%	10.2%	2
17097	59	3,220	5.5%	15.7%	2

Table 3.1 Percentile Ranking for Mesozones in Lake County, Illinois: *NAICS 31 to 33 (continued)*

County	Mesozone	Employment	Percentage of Employment	Cumulative Percent	Percentile Rank
17097	52	4,679	8.0%	23.7%	3
17097	61	6,741	11.5%	35.2%	4
17097	60	8,076	13.8%	49.0%	5
17097	56	8,778	15.0%	63.9%	7
17097	53	21,138	36.1%	100.0%	10

3.4 Supply Chain Formation

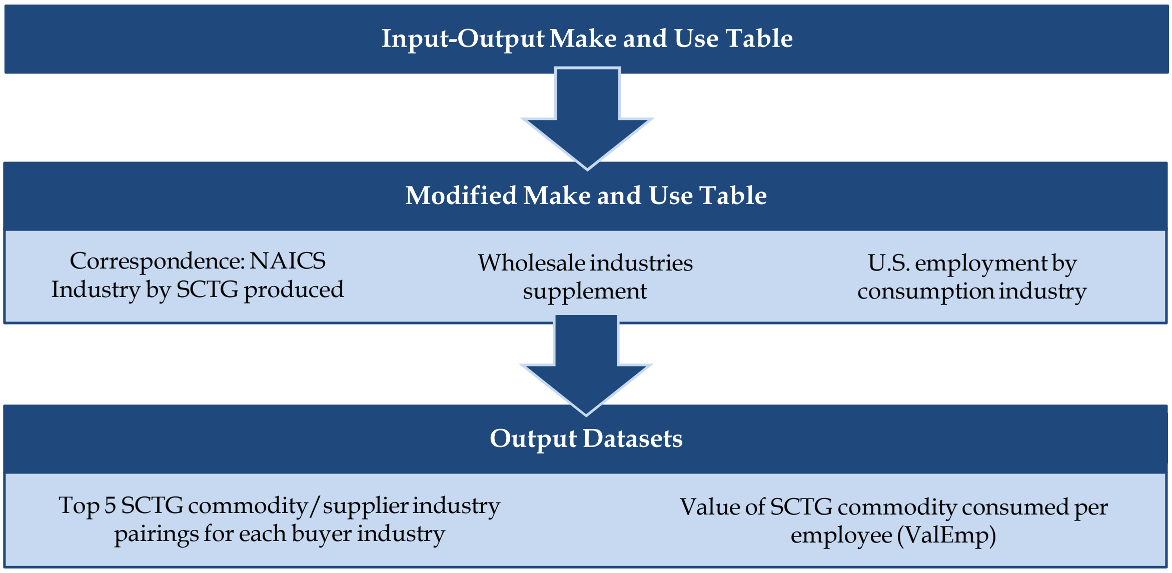
In the mesoscale model, shipper-receiver pairs are discrete decision-making units. This step evaluates characteristics of the firms from Step 1 and pairs each buyer or consumer firm with a supplier or producer firm.

This step involves the following subtasks. First, a choice set of suppliers is prepared for each buyer. Second, to create the choice set, suppliers are screened by industry class and type of commodity produced based on supply chain information from the Make and Use tables. Third, further screening is performed using reported flows from the FAF3 dataset. This ensures that suppliers are only selected from reported production regions for a specific commodity-attraction region combination. Finally, the buyers select a supplier to form the supply chain. High-level decision-making rules are used for the prototype.

3.4.1 The Supplier Choice Set

The supplier choice set preparation begins with the 2002 Benchmark Input-Output Make and Use table, which is used to identify the **top five supply industries** for each buyer firm in the mesoscale model. This table contains the amount of goods exchanged between industries by value. For example, grain farmers sell their products to bread manufacturers, cattle ranchers, flour millers, breweries, and other industries. The Make and Use table indicates how much of the total grain farm output goes to each of these industries. Figure 3.3 shows how the Make and Use table is invoked in the model stream.

Figure 3.3 Make and Use Table Processing



Focusing on the top five commodities (by value) that are used by each industry is a constraint that was imposed to overcome three practical concerns. First, future data collection efforts will probably have similar constraints. It would be difficult to collect information on input commodities and trading partners for an exhaustive list of commodities that are used by a particular buyer. Second, minimizing model run time during model development and testing was a key consideration. Finally, focusing on a handful of commodity trades for each industry makes the functioning of the prototype model more coherent to the user – a key consideration in this prototype development effort. This constraint can be removed for the full implementation.

The mesoscale model then uses simulation to evaluate whether a buyer works directly with a supplier or uses a wholesale intermediary to supply the necessary commodity. The supplier industry list for each buyer is modified and the Make and Use table is adjusted to allow this process to be streamlined with the direct shipper-receiver processing. Although this process is an imperfect fit within the overall behavioral framework of the mesoscale model, it allows the model design to better match the sampling plan used in the FAF3 framework, where flow data were sampled from both manufacturers and wholesalers.

3.4.2 *Identification of the Traded Commodity*

Next, the NAICS-SCTG correspondence is used to identify the type of SCTG commodity that is being exchanged between the buyer firm and each of its suppliers. In some cases – e.g., when a particular industry produces more than one type of commodity – random draws are used to make a final determination of the type of commodity that is being exchanged.

3.4.3 Refinement of Supplier Choice Set

The choice set of suppliers is further refined using reported flows and their O-D regions from the FAF3 dataset. These flows are reported using the two-digit SCTG classification codes. This screening ensures that suppliers are only selected from reported production regions for a specific commodity-attraction region combination.

3.4.4 Supplier Selection

Finally, each buyer selects one supplier for each input commodity that is used by the buyer firm. For the prototype, the following steps are implemented:

1. A sample of up to 10 suppliers from the full supplier choice set is drawn;

2. A random constant is generated for each potential supplier. This constant has two purposes. First, it is intended to represent unknown characteristics of the supplier. Second, it helps ensure that a variety of suppliers are selected (rather than the "best" supplier each time);

3. A utility-like value is calculated for each supplier based on the **Supplier Selection** parameters listed in Table 2.8. The random constant is added to this value; and

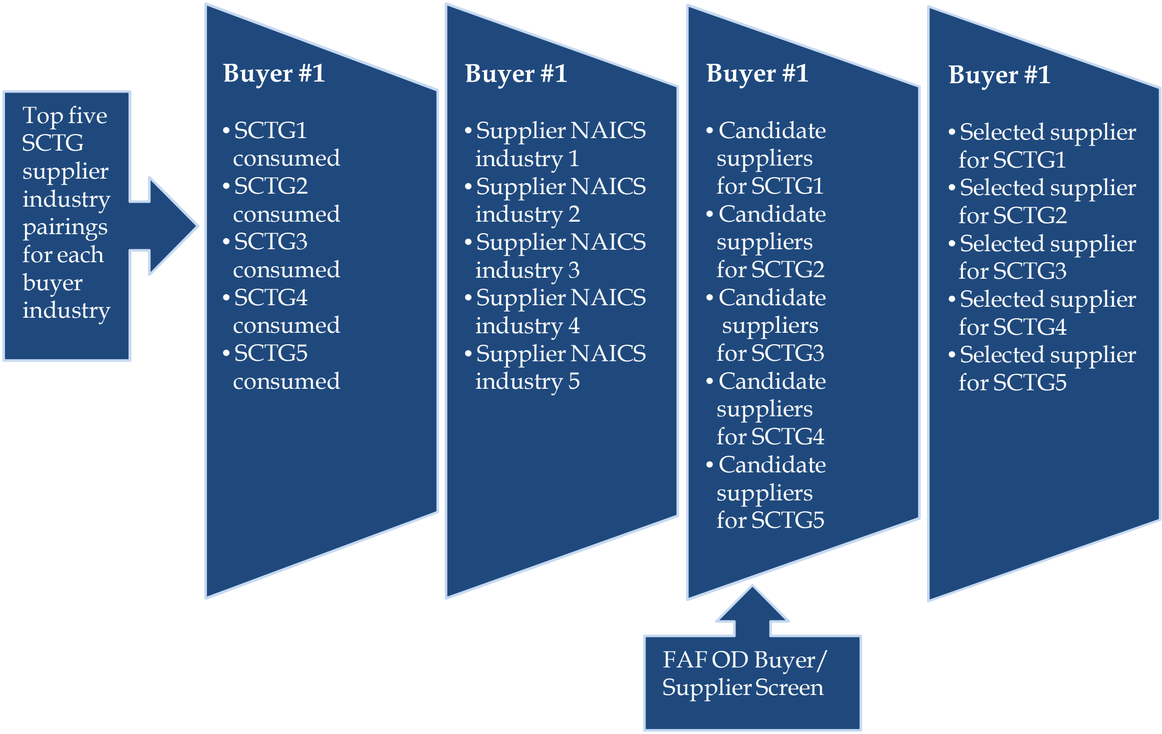
4. The supplier that is associated with the maximum value is selected.

Other methods of **Supplier Selection** can be used in the full mesoscale model implementation following a data collection and parameter calibration effort.

During supply chain formation, suppliers are identified as firm-types. Later in the model stream, these suppliers are identified as specific individual firms.

Figure 3.4 summarizes the **Supplier Selection** methodology. This figure illustrates the **Supplier Selection** process for a single buyer in the dataset.

Figure 3.4 Supplier Selection Flowchart



3.5 Apportionment of Flows to Shipper-Receiver Pairs

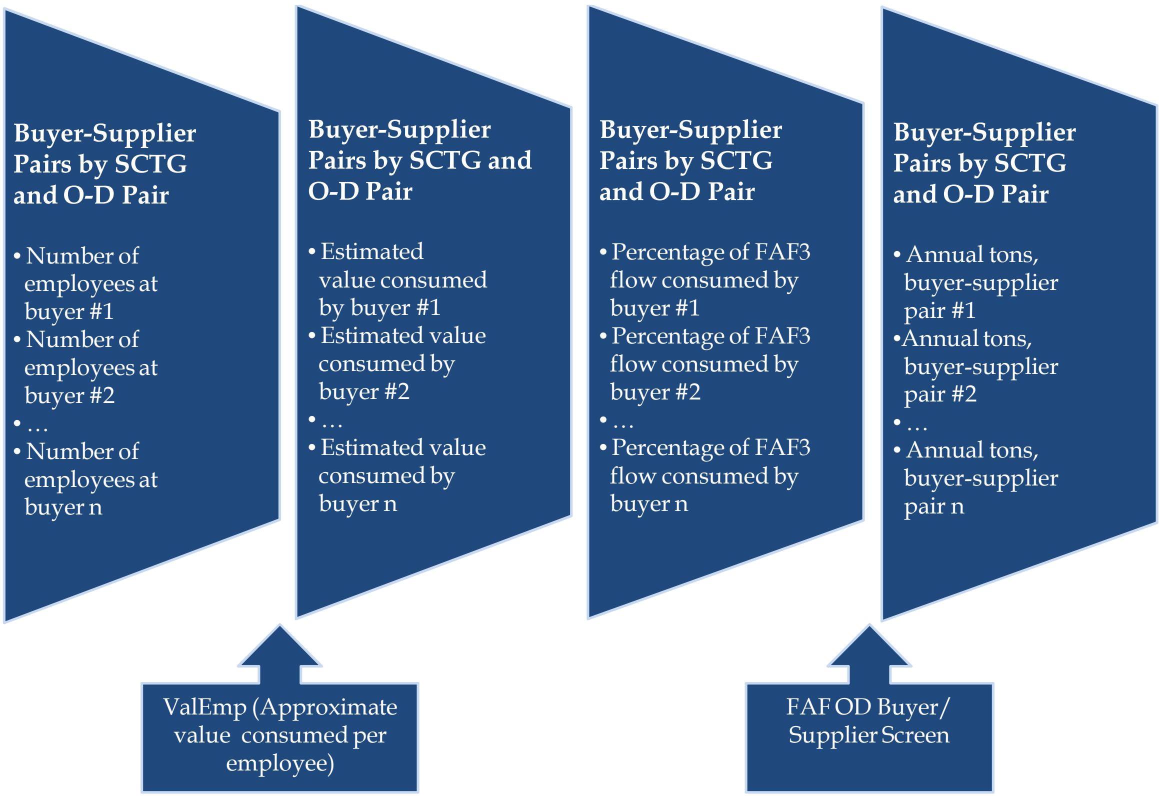
In this step, FAF3 flows for a given pair of origin and destination zones are apportioned to the individual shipper-receiver (or supplier-buyer) pairs that are associated with the same zone pair. The key inputs to this process are:

- The FAF3 flow data for the Chicago region (see Figure 3.2);
- The shipper-receiver pairs from the above step;
- The number of employees at the buyer firm; and
- The volume of goods consumed per employee, which is derived using total industry employment from CBP data and total value purchased by each industry in the Make and Use table (Figure 3.3).

The third and fourth inputs are used to calculate the percentage of flow by commodity that is consumed by each buyer. These percentages are multiplied by the FAF3 volumes to calculate total annual flow by commodity for each shipper-receiver pair.

Figure 3.5 illustrates how the flows are disaggregated from the FAF3 O-D pairs to individual shipper-receiver pairs.

Figure 3.5 Apportionment of Commodity Flows to Shipper-Receiver Pairs



3.6 Modeling of Transport and Logistics Path Choices

In this step, transport and logistics-related choices are evaluated for shipments between each shipper and receiver that form a supply chain. These choices are modeled jointly using carefully designed skim building and path formulation procedures. Evaluating the selected shipment frequency and the selected transport and logistics path simultaneously is a key element of the mesoscale framework.

The EMME/3 network and skimming procedures that were developed for the mesoscale model generate the following transportation path characteristics:

- Line-haul distance;
- Network access distance (distance from origin to the line-haul network, if applicable);

- Network egress distance (distance from the line-haul network to the final destination, if applicable);
- Logistics handling facility used in the CMAP region; and
- Logistics handling facility used outside of the CMAP region (represented generically).

The transport skims are read into SAS and further processed to include the following additional path characteristics:

- Mode;
- Shipment size;
- Inventory-related costs;
- Other miscellaneous costs such as stock-out and ordering costs; and
- Total annual transport and logistics costs.

Section 1.0 describes the individual components that make up total path costs. Following the data collection effort, some of these costs may need to be simplified for the fully calibrated mesoscale model implementation.

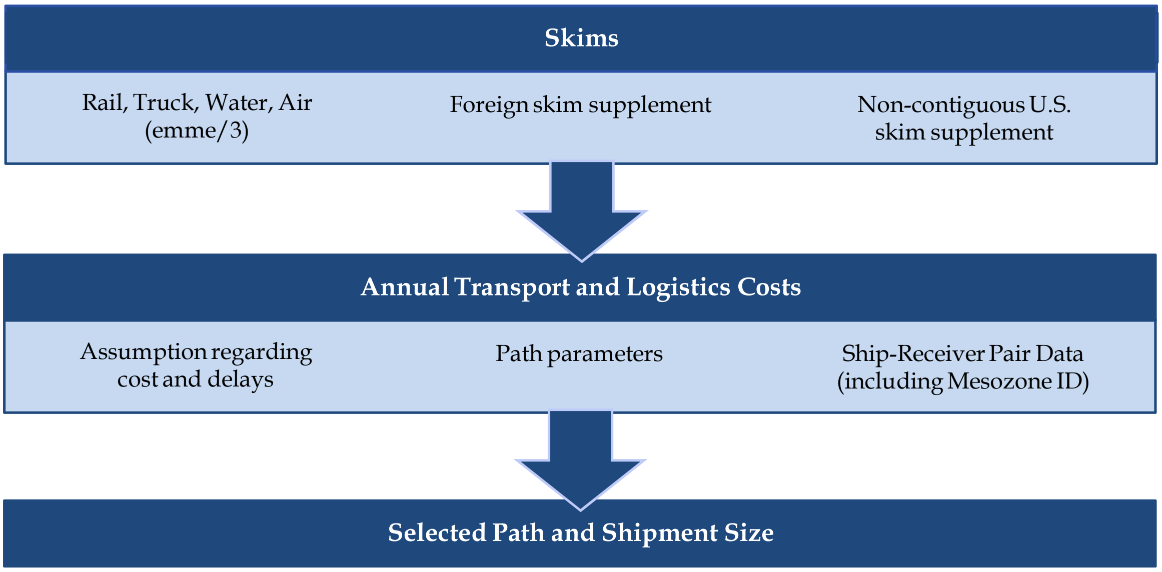
The path cost parameters that are used in the prototype are shown in Table 2.9. Calibration of these parameters will be important for fully implementing and applying the mesoscale model.

Total annual transport and logistics cost varies by shipment frequency. There are many possible reasons for this – e.g., carriers may offer volume discounts or shippers may apply a handling fee each time an order is placed. The prototype mesoscale model addresses this relationship by evaluating these decisions simultaneously.

Path Selection is performed in a way that is similar to the **Supplier Selection** process. A utility-like value is generated for each path, then the path with the maximum value is chosen. The utility-like value is the sum of a random constant is generated for each path and the combined "utility" of the other path components, which are calculated using the skims in combination with the parameters in Table 2.9.

This processing results in a dataset of shipper-receiver pairs that identifies the commodity being exchanged, the annual volume, shipment frequency, and details of the chosen path such as mode and location of regional logistics handling stops. Figure 3.6 shows the model processing steps for the **Path Selection** process.

Figure 3.6 Path Selection Process



■ 3.7 Postprocessing: Preparing Results for Assignment

Preparing the results for assignment involves breaking down individual shipments by trip leg, converting tons to vehicle trips, and generating empty truck trips. Because truck has the highest freight mode share in the CMAP region, and because the highway network is heavily funded by public dollars, modeling trucks will probably be important for the freight assignment.

For the prototype model, a rudimentary trip table preparation was performed to conduct reasonableness checks of the final results. CMAP can build upon this effort or develop a new procedure that performs a similar conversion of path-based shipments into individual trips. The basic steps are as follows:

- First, each goods movement can be separated into a maximum of three individual trip legs. Shipments that stop at one or two handling locations can be broken down into two trips with the handling facility (or facilities) becoming a new origin and destination point. Handling stops in the CMAP region are naturally of greater interest than stops in external areas.

- Second, tonnages are converted into vehicles for assignment. Widely accepted payload factors are available from VIUS. However, it is important to note that these factors were developed based on full and less than full trucks. Another approach could involve estimating density of goods using other sources and developing an alternate set of payload factors by commodity.

- Third, empty truck trips must be prepared. Most likely, these are trucks that are being repositioned for carrying the next load. Some percentage of commodity-carrying trucks will locate a load in the region for the backhaul portion of the trip. The remainder will probably need to travel to another region to pick up the next load. Empty truck trips can be developed based on these ideas.

Another decision to be made is what network to use for assigning the mesoscale freight vehicles. Options include:

- The passenger model network, enhanced with freight-related network attributes, can be used for a multiclass assignment. This would require a subarea extraction of a mesoscale assignment. Alternatively, the point of entry can be recorded during the mesoscale skimming process;

- The freight assignment can be performed on the mesoscale freight network with background traffic from the passenger model. However, both freight and passenger classes would ideally be modeled together to include the effects of congestion; or

- Salient features of the mesoscale network, such as the national portions of the network, can be stitched to the regional passenger model. The regional passenger model can be enhanced to accommodate additional freight network detail.

Ultimately, the mesoscale outputs also will need to be prepared for input to the microscale model. However, at this time the microscale specifications are not developed enough to make any recommendations regarding the mesoscale-microscale link.

4.0 Model Application

This section discusses the application of the model as follows. First, it documents tests that were conducted to ensure that the model functions as expected. Basic statistics such as FAF3 tons input to the model and tons that are output from the model are compared.

Second, this section presents preliminary statistics that are output from the draft prototype model. These preliminary statistics demonstrate the type of output that can someday be obtained from a fully calibrated mesoscale model **following the anticipated data collection[16] and model calibration effort**. It is important to emphasize that the prototype is a working demonstration that is intended to illustrate the usefulness of a fully calibrated model; it has limited use for immediate applications due to lack of calibration data.

Third, because the model eventually is intended to be used for analyzing a variety of freight-related questions, this section discusses potential ways in which the fully calibrated mesoscale model can be applied. Limitations of the model for application purposes also are discussed.

4.1. Model Testing

The objective of the model testing phase was to ensure that the model behaves as expected. This section presents the results of two high-level tests. Preliminary outputs that are shown later in this section also indicate that the model is functioning as expected. Additional tests were conducted during development to ensure that each step was working properly. These are summarized in Section 5.3.3.

In Table 4.1, the flow volumes by commodity are documented at critical junctures of the model:

1. The raw FAF3 input flows;
2. The processed FAF3 flows following apportionment to the CMAP region;
3. The flows following **Supplier Selection**;
4. The factored flows following **Supplier Selection** (the factoring process is described in the next section); and

[16] Section 2.0 describes the data that are necessary for calibrating the model.

5. The flows after **Path Selection**.

The flow totals for the first and second columns show that about half of the raw FAF3 input flows are maintained in the mesoscale model. This is a reasonable result considering that the raw input flows cover all of Wisconsin, nearly all of Illinois, and the entire northwest Indiana FAF3 zone. A comparison of the second and third columns shows that about two to three percent of flows are dropped during the **Supplier Selection** module. The reasons for this are explained in the next section. Commodity-specific factors are applied in the fourth stage to adjust these flows to match the input flows from the second stage. Finally, the flow totals following **Path Selection** are checked to ensure that they also match the flows from the second stage.

Table 4.1 Flow Totals[a] at Key Stages of the Model

	Tons (Year 2007)				
	FAF3		Model		
SCTG2	1. Includes All Five Overlapping FAF3 Zones	2. Customized to CMAP Region	3. After Supplier Selection (No Factors)	4. After Supplier Selection (with Factors)	5. After Path Selection
1	6,052,000	736,555	669,841	736,555	736,555
2	228,558,000	41,869,784	40,167,349	41,869,784	41,869,784
3	51,980,000	18,013,463	17,902,753	18,013,463	18,013,463
4	26,935,000	5,637,284	5,547,208	5,637,284	5,637,284
5	12,409,000	6,094,487	6,074,638	6,094,487	6,094,487
6	18,134,000	10,539,757	9,750,714	10,539,757	10,539,757
7	63,382,000	31,121,227	30,675,451	31,121,227	31,121,227
8	15,547,000	8,726,645	5,040,165	8,726,645	8,726,645
9	172000	107,406	107,207	107,406	107,406
10	7,477,000	6,080,536	5,935,642	6,080,536	6,080,536
11	40,214,000	25,661,092	25,081,054	25,661,092	25,661,092
12	154,388,000	65,421,552	65,379,328	65,421,552	65,421,552
13	46,627,000	36,343,777	35,548,931	36,343,777	36,343,777
14	24,607,000	9,234,766	7,650,347	9,234,766	9,234,766
15	155,693,000	37,058,030	34,442,661	37,058,030	37,058,030
16	327000	197,683	121,207	197,683	197,683
17	27,200,000	18,362,156	18,003,388	18,362,156	18,362,156
18	15,257,000	5,773,454	5,773,105	5,773,454	5,773,454
19	62,882,000	48,798,370	48,267,898	48,798,370	48,798,370

Table 4.1 Flow Totals[a] at Key Stages of the Model (continued)

	Tons (Year 2007)				
	FAF3		Model		
SCTG2	1. Includes All Five Overlapping FAF3 Zones	2. Customized to CMAP Region	3. After Supplier Selection (No Factors)	4. After Supplier Selection (with Factors)	5. After Path Selection
20	43,070,000	34,288,302	34,100,966	34,288,302	34,288,302
21	1,859,000	1,516,050	1,511,092	1,516,050	1,516,050
22	30,205,000	20,435,673	20,160,409	20,435,673	20,435,673
23	18,327,000	12,771,179	12,739,092	12,771,179	12,771,179
24	23,833,000	15,834,020	15,684,516	15,834,020	15,834,020
25	3,082,000	1,448,876	1,440,239	1,448,876	1,448,876
26	28,634,000	13,542,220	13,070,683	13,542,220	13,542,220
27	27,837,000	14,036,501	13,738,932	14,036,501	14,036,501
28	13,836,000	7,421,132	6,375,847	7,421,132	7,421,132
29	10,808,000	5,950,352	5,726,966	5,950,352	5,950,352
30	3,754,000	2,414,978	2,375,619	2,414,978	2,414,978
31	101,194,000	71,085,612	67,447,827	71,085,612	71,085,612
32	76,965,000	56,990,549	56,601,904	56,990,549	56,990,549
33	20,698,000	13,577,282	13,532,888	13,577,282	13,577,282
34	18,146,000	10,223,581	10,204,734	10,223,581	10,223,581
35	8,799,000	5,527,156	5,515,669	5,527,156	5,527,156
36	16,606,000	9,697,156	9,541,169	9,697,156	9,697,156
37	1,997,000	1,648,059	1,614,273	1,648,059	1,648,059
38	756000	568,530	565,808	568,530	568,530
39	4,570,000	2,856,351	2,740,069	2,856,351	2,856,351
40	12,210,000	8,432,175	8,323,618	8,432,175	8,432,175
41	96,025,000	69,911,210	68,156,661	69,911,210	69,911,210
Total	**1,521,052,000**	**755,954,969**	**733,307,869**	**755,954,969**	**755,954,969**

[a] Flow totals include flows that originate and/or terminate in the described geographic region. Other filtering (such as removing pipeline flows) that was described earlier in this report also applies to all of the values shown in this table, including the Stage 1 – FAF3 input totals.

Next, the results of the business location model are examined. This component is described in Section 3.3.2. The objective of this process was to assign each individual firm to a specific mesozone in the county within which the firm is located. The process uses percentile rankings to assign larger firms to mesozones with more employment and smaller firms to zones with any employment in their industry classes. The model prevents firms from being assigned to Mesozones with no employment in the firm's industry class.

Table 4.2 shows the results of the business location model for firms in NAICS classes 31 to 33 in Lake County, Illinois. In order to check the business location model results, total employment is summed up over all firms in each mesozone and compared (using percentages) to the total mesozone employment that was derived from subzone data. The comparisons in this table indicate that the model is functioning as expected: zones with more observed employment are assigned more employment overall than zones with less observed employment.

Table 4.2 Distribution of Modeled Versus Observed Employment in Lake County (IL) Mesozones: *NAICS 31 to 33*

Meso Zone	Percentile Rank within County (3=highest)	Percentage of Employment within County	
		Modeled	Observed
46	1	1%	0%
48	1	1%	0%
49	1	1%	0%
50	1	1%	1%
58	1	1%	1%
47	1	1%	1%
57	1	1%	2%
55	1	1%	2%
51	1	1%	2%
54	2	5%	2%
59	2	6%	5%
52	3	4%	8%
61	4	10%	12%
60	5	15%	14%
56	7	22%	15%
53	10	32%	36%
Total		100%	100%

Table 4.2 also points to some areas of potential improvement for this process. For example, modeled employment is somewhat higher or lower than observed employment in several instances. Procedures such as iterative looping could be introduced to improve the overall fit.

In addition, for the full implementation of the mesoscale model, other features can be added to the business location model to enhance its explanatory power. For example, business location could be a function of transportation accessibility and proximity to market.

■ 4.2 Preliminary Statistics

This section presents preliminary statistics from the uncalibrated prototype model. As described earlier, the objective of presenting these statistics is to illustrate some of the analytical capabilities of the mesoscale model. The model must be calibrated with observed or stated-preference data before it can be employed for planning purposes.

Table 4.3 shows the total number of buyer and supplier firms by number of employees that are created during **Firm Generation**. Placeholder data for foreign zones were generated by assuming that each foreign zone has businesses in each NAICS and size category.

Table 4.3 Total Buyer and Supplier Firms Generated by Firm Size and Location

	Location			
Number of Employees	**Domestic – CMAP Region**	**Domestic – Not in CMAP**	**Foreign**	**All Locations**
1 to 19	244,150	6,987,394	-	7,231,544
20 to 99	34,276	940,190	-	974,466
100 to 249	5,793	136,289	-	142,082
250 to 499	1,370	35,447	-	36,817
500 to 999	549	13,931	-	14,480
1,000 to 2,499	245	7,052	3,032	10,329
2,500 to 4,999	45	2,744	3,032	5,821
Over 5,000	19	2,065	3,032	5,116
All Firms	286,447	8,125,112	9,096	8,420,655

Tables 4.4 and 4.5 show the number of supplier and buyer firms that are used in the mesoscale model. Table 4.4 shows the number of firms (or firm-types) by location and Table 4.5 focuses on firm size. The number of firms that is input into **Supplier Selection** is a subset of the universe of firms shown in Table 4.3. As described in Section 3, this happens because all CMAP agents are kept in the model but only a sample of agents

outside of the CMAP region are kept. The buyers in Tables 4.4 and 4.5 are maintained during the **Supplier Selection** process.

Furthermore, while all agents in the CMAP region are maintained, the total number of CMAP-based agents in Table 4.4 is less than the number of CMAP-based agents in Table 4.3. This happens for the following reason. If none of the top five supply industries for a given buyer are identified as industries that produce an SCTG commodity, then the buyer industry effectively is removed from the model.

For example, the top five "supply" industries for buyers that are classified as Input-Output NAICS 813B00 (Civic; social; professional and similar organizations) are funds, trusts and securities. These are really monetary transactions rather than freight trades. In reality, such buyers trade with suppliers of actual commodities such as printing and apparel commodities. As described in Section 3.4.1, modeling an exhaustive list of supplied commodities for each buyer would ensure that these transactions are included; however, doing so may be undesirable for other reasons.

Table 4.4 Number of Firms or Firm-Types (by Location) Input to Supplier Selection

	Location			
Producer/Consumer Category	**Domestic – CMAP Region**	**Domestic – Not in CMAP**	**Foreign**	**All Locations**
Buyer Firms	258,189	423,225	9,096	690,510
Supplier Firm-Types	4,616	28,218	6,853	39,687

Table 4.5 Number of Firms or Firm-Types (by Size) Input to Supplier Selection

Number of Employees	Buyer Firms	Supplier Firm-Types
1 to 19	538,187	9,606
20 to 99	87,750	10,761
100 to 249	20,946	6,603
250 to 499	8,472	3,302
500 to 999	14,047	1,596
1,000 to 2,499	10,196	3,046
2,500 to 4,999	5,800	2,428
Over 5,000	5,112	2,345
All Modeled Agents	690,510	39,687

If all buyers that are input into **Supplier Selection** are matched with a supplier, than the number of shipper-receiver pairs following **Supplier Selection** would be equal to the total number of buyer firms that are input into the process. However, a small percentage (approximately 0.5 percent) of inputted buyers are not matched with a supplier. That is, some buyer firms that are input into **Supplier Selection** do not form a supply chain. This occurs for the following reasons:

1. Observed flows in the FAF3 dataset are used to filter out firms that are located in regions where no trade has been reported with partners in the Chicago region. For example, if the FAF3 dataset indicates that there are no petroleum shipments from the CMAP region to Hawaii, then petroleum buyers in Hawaii will not form any supply chains with CMAP area suppliers;

2. The supplier choice set could lack suppliers of the correct "make" industry in a given FAF3 production region where flows reportedly originated. That is, the buyer may search for suppliers of industry X but not locate any in a production region; or

3. Buyers might not select each and every supplier in the supplier choice set during the **Supplier Selection** phase. This can happen when the ratio of buyers to suppliers is relatively low. Suppliers that are not chosen disappear from the model stream.

If it so happens that no suppliers are selected from a production region where observed flows originate, then the related flows are not apportioned during the **Flow Apportionment** stage. This happens to about two to three percent of flows in the model. There are several potential ways to address this, including the following:

- External zones can be made greater in size, resulting in fewer supplier firm-types for the supplier choice set. However, this would require a less detailed network, which would be a disadvantage;

- The model can be constrained to produce a shipper-receiver pair for each reported origin-destination pair. However, this would diminish the power of the model's behavioral analysis, which is undesirable; or

- The existing flows can be factored upwards to account for the missing flows. This option was implemented for the prototype model.

Table 4.6 shows the distribution of the buyer and supplier firms by number of employees following **Supplier Selection**.

Table 4.6 Distribution of Paired Buyers and Suppliers by Size

Buyer Size (Number of Employees)	Supplier Size (Number of Employees)							
	1 to 19	20 to 99	100 to 249	250 to 499	500 to 999	1,000 to 2,499	2,500 to 4,999	Over 5,000
1 to 19	980,932	646,369	266,111	154,272	254,260	129,161	61,080	50,578
20 to 99	158,797	107,538	48,324	23,568	38,666	19,452	9,983	8,276
100 to 249	17,471	10,042	37,438	16,654	8,313	3,990	1,953	1,406
250 to 499	7,423	4,250	15,515	6,707	3,097	1,194	489	390
500 to 999	18,569	11,697	17,962	9,252	4,781	1,084	290	364
1,000 to 2,499	14,358	8,725	12,980	6,173	2,877	603	175	220
2,500 to 4,999	8,535	5,042	7,061	3,188	1,290	259	106	85
Over 5,000	7,366	4,447	6,284	2,747	1071	226	83	53

Table 4.7 shows the mean great circle distances between the buyer and the selected supplier by commodity following the **Supplier Selection** module.

Table 4.7 Great Circle Distance between Buyers and Suppliers

SCTG	GCD, Supplier to Buyer	
	Mean	N (Pairs)
1 – Live animals/fish	1,060	6,242
2 – Cereal grains	569	32,539
3 – Other agriculture products	709	80,943
4 – Animal feed	1,004	87,737
5 – Meat/seafood	612	137,127
6 – Milled grain products	802	7,067
7 – Other foodstuffs	873	190,967
8 – Alcoholic beverages	960	478
9 – Tobacco products	985	79,129
10 – Building stone	848	1,304
11 – Natural sands	1,185	898
12 – Gravel	1,284	6,114

Table 4.7 Great Circle Distance between Buyers and Suppliers (continued)

	GCD, Supplier to Buyer	
SCTG	Mean	N (Pairs)
13 – Nonmetallic minerals	1,723	3,126
14 – Metallic ores	2,174	2,851
15 – Coal	1,428	1,271
16 – Crude petroleum	624	818
17 – Gasoline	120	24,929
18 – Fuel oils	461	23,611
19 – Coal-n.e.c.	817	101,328
20 – Basic chemicals	940	123,189
21 – Pharmaceuticals	761	91,179
22 – Fertilizers	977	21,445
23 – Chemical products	714	311,172
24 – Plastics/rubber	819	231,441
25 – Logs	1,057	17,264
26 – Wood products	877	99,236
27 – Newsprint/paper	780	81,142
28 – Paper articles	798	39,012
29 – Printed products	565	333,599
30 – Textiles/leather	1,128	32,044
31 – Nonmetal mineral products	1,291	7,830
32 – Base metals	1,010	48,697
33 – Articles-base metal	904	95,793
34 – Machinery	762	202,292
35 – Electronics	777	414,102
36 – Motorized vehicles	687	124,642
37 – Transport equipment	796	14,924
38 – Precision instruments	694	102,474
39 – Furniture	835	15,865
40 – Miscellaneous manufactured products	990	41,942
41 – Waste/scrap	645	13,889

Table 4.8 shows the distribution of the inverse of shipment frequency by the number of shipper-receiver pairs. For the prototype, four shipment frequencies were tested in the **Path Selection** calculations: once every two, four, six, and eight weeks. This can be expanded for the full implementation.

Table 4.8 Distribution of the Number of Weeks between Shipments

Number of Weeks between Shipments	Percentage of Shipper-Receiver Pairs
2	27%
4	12%
6	8%
8	53%

Table 4.9 lists the modeled percentage of annual tonnages by commodity type that use each path. The path is selected by the shipper-receiver pairs during the **Path Selection** phase. The figures in this table represent the total flows between all shipper-receiver pairs that trade a particular type of commodity such as bulk goods. Paths that were not selected by any agents are not shown.

Table 4.9 Percentage of Annual Volume Using Each Type of Path

	Percentage of Annual Tons by Type of Goods			
Path Description	Live Fish/ Animals	Bulk Natural Resource (BNR)	Intermediate Processed Goods (IPG)/Other	Finished Goods (FG)
01 – Water using Port 145	0.0%	6.4%	0.9%	0.3%
02 – Water using Port 146	0.0%	4.2%	1.4%	0.3%
03 – Carload Direct	0.2%	1.0%	1.0%	3.0%
04 – FTL (ExtDray)-Carload remainder	0.1%	0.0%	0.2%	0.1%
05 – Carload-FTL with stop at 147	6.4%	4.7%	1.5%	1.5%
06 – Carload-FTL with stop at 148	0.3%	4.2%	0.9%	0.9%
07 – Carload-FTL with stop at 149	41.3%	6.2%	3.7%	3.7%
08 – Carload-FTL with stop at 150	36.5%	34.0%	46.5%	26.3%
13 – IMX Direct	0.2%	0.0%	0.1%	0.5%
14 – FTL (ExtDray)-IMX remainder	0.0%	0.0%	0.0%	0.1%
15 – IMX-LTL with stop at 147	0.0%	0.0%	0.1%	0.1%

Table 4.9 Percentage of Annual Volume Using Each Type of Path (continued)

	Percentage of Annual Tons by Type of Goods			
Path Description	**Live Fish/ Animals**	**Bulk Natural Resource (BNR)**	**Intermediate Processed Goods (IPG)**	**Finished Goods (FG)**
16 – IMX-LTL with stop at 148	0.0%	0.0%	0.2%	0.3%
17 – IMX-LTL with stop at 149	0.0%	0.0%	0.1%	0.1%
18 – IMX-LTL with stop at 150	0.4%	0.2%	4.1%	4.2%
20 – IMX-FTL with stop at 148	0.0%	0.0%	0.0%	0.3%
21 – IMX-FTL with stop at 149	0.0%	0.0%	0.0%	0.1%
22 – IMX-FTL with stop at 150	0.0%	0.0%	0.2%	5.4%
31 – FTL Direct	4.8%	37.8%	32.9%	41.4%
32 – FTL-LTL with stop at 133	0.3%	0.0%	0.0%	0.3%
33 – FTL-LTL with stop at 134	0.3%	0.0%	0.0%	0.2%
34 – FTL-LTL with stop at 135	0.6%	0.0%	0.1%	0.4%
35 – FTL-LTL with stop at 136	0.1%	0.0%	0.2%	0.5%
36 – FTL-LTL with stop at 137	0.0%	0.0%	0.3%	1.0%
37 – FTL-LTL with stop at 138	0.3%	0.0%	0.1%	0.7%
38 – FTL-LTL with stop at 139	0.5%	0.0%	0.2%	0.8%
39 – LTL-FTL-LTL with stop at 133	0.0%	0.0%	0.0%	0.1%
40 – LTL-FTL-LTL with stop at 134	0.0%	0.0%	0.0%	0.1%
41 – LTL-FTL-LTL with stop at 135	0.0%	0.0%	0.0%	0.1%
42 – LTL-FTL-LTL with stop at 136	0.0%	0.0%	0.0%	0.1%
43 – LTL-FTL-LTL with stop at 137	0.0%	0.0%	0.0%	0.2%
44 – LTL-FTL-LTL with stop at 138	0.0%	0.0%	0.0%	0.3%
45 – LTL-FTL-LTL with stop at 139	0.0%	0.0%	0.0%	0.2%
46 – LTL Direct	7.6%	1.5%	5.1%	6.4%
47 – Air using Airport 141	0.0%	0.0%	0.0%	0.0%
48 – Air using Airport 142	0.0%	0.0%	0.0%	0.0%
49 – Air using Airport 143	0.0%	0.0%	0.0%	0.0%
50 – Air using Airport 144	0.0%	0.0%	0.0%	0.0%
52 – Int'l. Water, No Transload, 40' Direct	0.0%	0.0%	0.0%	0.0%
53 – Int'l Water, Transload, 53 feet FTL	0.0%	0.0%	0.0%	0.0%
Total	**100.0%**	**100.0%**	**100.0%**	**100.0%**

Table 4.10 shows the total number of buyer-supplier pairs following **Path Selection** that would use a dray truck to transport shipments to or from a logistics node in the Chicago region. Dray trips that occur in external zones are not included in this table.

Table 4.10 Number of Buyer-Supplier Pairs Using Each Logistics Node in the Chicago Region

Logistics Node ID	Number of Shipper-Receiver Pairs
Truck terminal in Kenosha	80,270
Truck terminal in Rockford	147,033
Truck terminal near DuPage/Kane border	257,607
Truck terminal in Will County	301,113
Truck terminal in Hammond/Gary area	477,557
Truck terminal near I-294/290 junction	314,034
Truck terminal in Central Chicago	360,630
O'Hare	5,769
Midway	3,911
Gary	792
Milwaukee	1,652
Illinois International Port	1,342
Indiana Harbor	473
Rockford-area rail yards	4,181
Global III-Rochelle	8,274
Logistics Park – Elwood	4,311
Central Chicago rail yards	150,605
All logistics nodes	2,119,554

In conclusion, this section presented preliminary statistics from the prototype model as examples of the types of output that can be produced using a fully calibrated model. These statistics include a variety of information on the types of agents, shipments, and paths that are modeled using the mesoscale model. The next section provides a review of potential planning and policy applications of the model.

4.3 Applications: Modeling of Policy and Project Alternatives

The mesoscale freight model framework provides a powerful and flexible framework for goods movement analysis in the Chicago region. The macroscale and microscale models, when developed, will generate additional analytical capability.

Like other agent-based models, the mesoscale model is most powerful for analyzing the collective impact of individual choices across the modeled agents. For example, following data collection and a subsequent model calibration effort, the fully implemented model can be used to estimate mode choice under different fuel price or congestion scenarios. Table 4.11 describes some examples of potential model applications and lists the types of choices being made in the model as well as suggested ways of analyzing the data. These examples include more traditional analyses such as mode choice and VMT as well as more innovative types of analysis such as evaluation of industry trading partners or a comparative analysis of the attractiveness of different logistics facilities.

Table 4.11 Examples of Potential Model Applications

Description of Analysis	Choices	Suggested Level of Analysis
Mode choice	Truckload, less-than-truckload, rail carload, rail IMX, air, water	Summarize by commodity type, distance, county, or other relatively broad category
Type of logistics facility used	Rail terminal/intermodal yard, truck terminal, water port, airport	Compare tons of freight passing through different types of facilities
Location of selected facility	Variable	Compare tons of freight passing through specific facilities
Supply chain analysis	Supplier selection	Evaluate trading partners by type of industry, distance, and business size
Vehicle-miles traveled (VMT)	Path selection	Summarize by commodity type, mode, county, or other relatively broad category
Ton-miles traveled	Supplier selection	Summarize at county level
Freight volumes by rail line at regional entry/exit points	Class I railroads	Summarize by carrier, external region, and other broad categories
Freight volumes by interstate at regional entry/exit points	Interstate highways	Summarize by highway

Following the anticipated data collection, calibration, and validation effort, the mesoscale model also can be used to evaluate many of the specific freight-related recommendations that were made in CMAP's 2010 ***Regional Freight System Planning Recommendations Study.*** For the study, recommendations in the broad categories of infrastructure, policy, and operations were qualitatively evaluated based on system performance measures of accessibility, economic development, mobility and safety.

The remainder of this section presents Tables 4.12 through 4.14 that discuss if and how the mesoscale (or macroscale/microscale) model can be used to address actual infrastructure, policy and operations scenarios that are based on the plan recommendations. Additional scenarios also are included in light of recent developments in U.S. trade and the economy.

The prototype mesoscale model is well equipped to conduct many of these infrastructure and policy scenarios. Its applicability for some policy and most freight operations scenarios, however, is more limited due to the more macroscopic or microscopic nature of the scenarios. For these cases, the macroscale model may provide answers at the appropriate geographical and temporal resolution. In the future, integration of the mesoscale planning tool with the macroscale model, which will provide insights into economic supply and demand, and with the microscale model, which will simulate freight operations, should be considered.

Table 4.12 Potential Scenarios for Testing the CMAP Freight Prototype Mesoscale Model: *Potential Infrastructure Scenarios*

Description of Scenario	Impact Area	Comment on Scenario Testing Feasibility Using Mesoscale Model
Construct a new intermodal rail terminal.	Global	Add a logistics node intermodal terminal node with a location somewhere in the Chicago Region served by rail. Connect the new node to rail and to local highway network.
Construct a new major truck distribution center in a suburban location with lower transfer costs than other Chicago DCs.	Global	Add a Distribution Center (DC) node connected to both the long-haul trucking and local truck highway networks. Make the transfer costs at the DC 50 percent of other DC costs in the Chicago region.
Construct a new industry with considerable transportation and logistics expenditure.	Global	Should be modeled primarily in macroscale model. Additional mesoscale analyses include an increase in the industry employment in a township mesoscale zone and an increase in the commodity tonnage related to that industry in the FAF.
Improve draft near ports.	Global	Port/terminal capacity currently is not modeled. However, this could be introduced by increasing the utility of water mode at the port(s) of interest. Survey data would be needed for calibration.
Upgrade freight corridor or Lower maintenance costs of freight infrastructure. (highways/rail)	Local	Most appropriate for microscale model. However, impacts can be approximated using mesoscale model by modifying operation costs on the network links of freight corridor or trade route. (highways/rail)
Capacity improvement or Landside access improvement or Bottlenecks elimination projects bundle. (highways/rail)	Local	Most appropriate for microscale model. However, to use the mesoscale model, network links, link capacities, and operation costs would need to be added or modified.
Construct dedicated truck lanes.	Local	Most appropriate for microscale model. However, to use the mesoscale model, would modify network links and/or link capacities, and place use restriction.

Table 4.12 Potential Scenarios for Testing the CMAP Freight Prototype Model (continued)
Potential Infrastructure Scenarios

Description of Scenario	Impact Area	Comment on Scenario Testing Feasibility Using Mesoscale Model
Build rail-highway grade separation projects bundle (such as CREATE's 25).	Local	Most appropriate for microscale model. However, to use the mesoscale model, would need to modify network links.
Implement safety enhancing technologies, including connected vehicles.	Local	Most appropriate for microscale model. However, to use the mesoscale model, would need to modify link capacity accounting for shorter headways on safety technology designated network links.
Implement freight infrastructure development programs, including CREATE.	Local	Mesoscale model can assess some impacts, but the microscale model would be most appropriate for understanding detailed operational benefits.
Construct truck parking facilities.	Local	Most appropriate for microscale model. Cannot be evaluated using mesoscale model.

Table 4.13 Potential Policy Scenarios

Description of Scenario	Impact Area	Comment on Scenario Testing Feasibility
Divert Asia shipments from LA to Seattle.	Global	Adjust the O&Ds in the FAF database such that tonnage is shifted from the FAF LA region to the FAF Seattle region. This is accomplished by a matrix addition and subtraction. Modify foreign entry points in mesoscale model.
Divert Asia shipments from LA to Houston (Panama Canal shift).	Global	Adjust the O&Ds in the FAF database such that tonnage is shifted from the FAF LA region to the FAF Seattle region. This is accomplished by a matrix addition and subtraction. Modify foreign entry points in mesoscale model.
Increase fuel prices.	Global	Modify operation costs for affected modes/vehicle technologies.
Increase transportation service costs.	Global/ Local	Modify operation costs for affected modes/vehicle technologies.
Revise length, size and weight regulations (regionally/nationally).	Global/ Local	Adjust payload factors used in transport and logistics path selection. Add/ modify operation costs of altered mode/vehicle technology.
Implement user fees concepts, such as VMT tax, toll, congestion pricing, etc. (regionally/nationally).	Global/ Local	Identify affected regions and increase costs for modes/vehicle technologies or links.
Growth in shortline and regional freight railroad.	Local	Add/modify economies of scale-related costs on links likely to be operated by shortline and regional freight railroad. Increase zonal employment in rail transportation and the dependent sectors.
Rationalize existing industries in light of economic recovery.	Local	Reduce zonal employment, reduce flows to/from that zone in the FAF OD table.
Increase clustering or concentration of businesses of core industries.	Local	Modify zonal employment of core industries.
Redevelop of brownfield lands for freight use.	Local	Cannot be fully evaluated using the mesoscopic model. Add/modify logistics node and node operational costs.

Table 4.13 Potential Policy Scenarios (continued)

Description of Scenario	Impact Area	Comment on Scenario Testing Feasibility
Supply information about freight infrastructure to manufacturers, distributors, service providers, freight handlers and other users in the Chicago region.	Local	Cannot be evaluated using the mesoscopic model.
Effect of driver shortages or shortages of skilled personnel.	Local	Increase cost for trucking modes.
Provide logistics-related training and educational opportunities.	Local	Cannot be evaluated using the mesoscopic model.
Innovative funding of freight projects.	Local	Cannot be evaluated using the mesoscopic model.
Create CREATE counterpart for trucking industry.	Local	Cannot be evaluated using the mesoscopic model.
Establish a freight authority for guiding investments and protecting the public interest.	Local	Cannot be evaluated using the mesoscopic model.
Preserve freight facilities from encroachments.	Local	Cannot be evaluated using the mesoscopic model.
Accommodate freight in ways conducive to passenger traffic.	Local	Cannot be evaluated using the mesoscopic model.
Implement freight safety programs, such as truck-lane restriction.	Local	Cannot be evaluated using the mesoscopic model.

Table 4.14 Potential Operations Scenarios

Description of Scenario	Impact Area	Comment on Scenario Testing Feasibility
Reduce costs at DCs in Chicago relative to national DCs.	Local	See if logistics chain shares shift from "direct to receiver" to "through Chicago DC centers."
Improve efficiency of terminal operations.	Local	Most appropriate for microscale model, but can use mesoscale model by modifying node operational costs due to delay (total of security screening, loading/unloading, parking, waiting in queue, switching, etc.).
Implement emission reduction technologies.	Local	Cannot be fully evaluated using the mesoscopic model. Utilize market penetration for ODCs, modify mode/vehicle technology or link operation costs.
Reduce empty container and bobtail moves.	Local	Lower costs for paths that use containers.
Reduce logistics costs by industries.	Local	Lower costs for paths that use containers.
Increase hours of service.	Local	Lower costs for paths that use containers.
Implement freight transportation time of day restriction.	Local	Most appropriate for microscale model; could modify cost and travel times in mesoscale model.
Implement measures for control of freight-related noise.	Local	Cannot be evaluated using the mesoscopic model.
Perform travel demand management for more efficient freight movements.	Local	Most appropriate for microscale model; could modify cost and travel times in mesoscale model.
Construct freight management center with technologies for freight tracking, performance measurement, and data-sharing.	Local	Most appropriate for microscale model; could modify cost and travel times in mesoscale model.
Improve supply chain visibility.	Local	Cannot be evaluated using the mesoscopic model.
Implement other regional ITS technologies.	Local	Cannot be evaluated using the mesoscopic model.

5.0 User's Guide

This section provides direction to the user in setting up and running the model. These instructions are intended to accompany the electronic data files that are provided with the report.

5.1 Hardware and Software Requirements

The model uses EMME/3 and SAS. The user's version of SAS must have the SAS/Access module for reading in *.dbf and Access database files. Due to the sizes of the model files, a powerful computer is recommend for running the SAS portion of the model. For the model development effort, a modern machine with eight 3.20-GHz processors was used. This resulted in run times of approximately three to four hours. If slow performance is detected, rebooting the machine may be helpful.

5.2 Setting up and Running the Model

The user can upload the files from the disk to a directory in his/her computing environment. The EMME/3 command files will automatically work with the EMME/3 data files in the same folder. The directory names that are associated with the SAS input data files will need to be renamed to reflect the directory to which they have been uploaded.

The EMME/3 code generates skims for the model. To run the script, the user should open the Prompt Window in EMME/3 then type in "~<MesoSkims.mac," which launches the macro command file that performs the skimming procedures. This file reads in the network, modes, rail routes, and other items needed to generate skims. The user will be asked if s/he wants to erase existing files in the output directory. The user generally should answer "Y" for yes, but should back up the existing files first if they will be needed later.

The skims that are output from the EMME/3 run are read by the SAS scripts. Therefore, the EMME/3 skim output files must be located in a directory that can be identified by the SAS program. This is a simple matter of either copying and pasting the files into the pre-specified directory or renaming the directory names in the SAS script.

The SAS programs are run by double-clicking on the program name then directing SAS to run (for example, using the menu, hitting F8, or clicking on the Run menu shortcut). The ***Data Inputs*** program reads in all of the data inputs and prepares them for using in the mesoscale model stream. The ***Meso Model*** program processes these data in accordance with the framework that is described in this report. A description of each of these steps is provided in Appendix B.

5.3 Tips and Troubleshooting

5.3.1 Emme/3 Skimming Tips

The existing EMME/3 network serves to demonstrate that the prototype model is functioning properly. However, as described in Section 2.2.1, a more detailed network can be developed for the full implementation of the mesoscale model. Following this development, the EMME/3 files will need to be updated accordingly, possibly replicating the same processes that are currently used.

The SAS files will also need to be updated when the EMME/3 network is expanded. For example, the ***Data Inputs*** command file contains macro variables in Step 1 that describe node ranges of important network features. Other items, such as formats (also in Step 1), also will need to be updated.

5.3.2 SAS Command File Tips

The SAS code is organized into sections, or steps, that make it easier for the user to understand the model stream. Additionally, the SAS code contains comments that describe the purpose of each step and its sub-tasks.

Each step that is listed in the code corresponds a step number in the Quick Reference guide (Appendix B). The Quick Reference guide documents the main input, intermediate, and output files that are featured throughout the model stream. Important files are color coded throughout the guide so that the reader can easily see where key files are being used or produced. It is also worth noting that important files are saved in a permanent SAS library location. Less critical files are housed during the session in the Work library.

Within the SAS program files, the model is organized into the following steps:

- Data Inputs.SAS:

  - Step 1. Define macro variables, libraries, filenames, file paths and variable formats;
  - Step 2. Read in and format datasets;
  - Step 3. Prepare employment data for CBPZONE system;
  - Step 4. Prepare data for various Mesoscale modules;
  - Step 5. Apportion FAF flows to CBPZONE system; and
  - Step 6. Prepare skims (GCD and EMME/3).

- Meso Model.SAS:
  - Step 1. Define libraries, filenames, file paths and variable formats;
  - Step 2. Firm generation;
  - Step 3. Supplier selection;
  - Step 4. FAF3 flow apportionment;
  - Step 5. Business location assignment;
  - Step 6. Path selection;
  - Step 7. Prepare trip table; and
  - Step 8. Apply formats & output data.

The user can easily search the code for a particular step using Ctl+F and typing in "Step 1" (or "Step 2", etc.).

5.3.3 Troubleshooting

The model has been thoroughly tested to ensure that is free of bugs. In addition, it has been tested to ensure that the files, once installed, run properly on different machines.

When making changes to the model, the user would benefit from conducting QA/QC checks that are similar in nature to those that are described in Section 4.1. Furthermore, as discussed in Section 4.1, a number of other checks were conducted that were not documented in the model. Examples of useful QA/QC checks are:

- Checking the SAS log file for warnings and errors and investigating potential sources of any unexpected messages;
- Making sure that the input totals match the output totals at various stages of the model;
- Searching the code (for example, using Ctl+F) to locate variables that seem to be causing issues; and
- Examining input and output data for a select set of individual observations.

Key features of the prototype model (such as **Supplier Selection** and **Path Selection**) will be improved upon during the full implementation when the necessary data become available. As a result, conducting extensive "troubleshooting" in such areas is not recommended at this stage of development. This effort falls under the domain of model calibration and validation, which is expected to be conducted using the data that are described in Section 2.2.

5.4 Evaluating Results

5.4.1 Model Outputs

The SAS program has been automated to generate a specific set of outputs. Other meaningful outputs can be created by the user. The user is expected to have some experience in SAS that will enable him/her to customize outputs for other analysis.

5.4.2 Alternatives Analysis

The model was designed with alternatives analysis in mind. The user can create network-based alternatives by modifying the EMME/3 network. For example, a new regional airport can be modeled. Some network changes, such as modifying the rail carrier routes, will require a corresponding modification of the EMME/3 input/output files. Additionally, some network changes must be accompanied by changes to the SAS command files. For example, the begin/end node numbers that define the node ranges for specific categories of facilities must be updated in the macro variable list of the ***Data Inputs*** file and format names and descriptions must be changed.

Appendix A: Development of Agricultural Employment Data

Employment for the agriculture sector of the U.S. was estimated using the FAF3 2007 data. The agricultural production value by FAF zone along with an assumed dollar output per employee were used to estimate employment in the FAF regions.

First, a crosswalk was created between the SCTG2 codes and the six-digit North American Industry Classification System (NAICS6) codes. Table A.1 shows the crosswalk used.

Table A.1 Industry-Commodity Crosswalk

NAICS6	Industry	SCTG2	Commodity
1111B0	Grain farming	02	Cereal grains
1111A0	Oilseed farming	03	Other agricultural products
111200	Vegetable and melon farming	03	Other agricultural products
1113A0	Fruit farming	03	Other agricultural products
111335	Tree nut farming	03	Other agricultural products
111400	Greenhouse, nursery, and floriculture production	03	Other agricultural products
111910	Tobacco farming	03	Other agricultural products
111920	Cotton farming	03	Other agricultural products
1119A0	Sugarcane and sugar beet farming	03	Other agricultural products
1119B0	All other crop farming	03	Other agricultural products
1121A0	Cattle ranching and farming	01	Live animals and live fish
112120	Dairy cattle and milk production	01	Live animals and live fish
112A00	Animal production, except cattle and poultry and eggs	01	Live animals and live fish
112300	Poultry and egg production	01	Live animals and live fish

Second, production values for live animals and live fish (SCTG 01), cereal grains (SCTG 02), and other agricultural products (SCTG 03) were extracted by FAF zone. However, the SCTG2 commodities shown in Table A.1 match more than one NAICS code. Ratios had to be prepared for distributing the SCTG2 FAF data among the NAICS industries. For SCTG 02 – cereal grains – the ratio is one because this commodity group only matches one industry code.

However, for SCTG 01 and SCTG 03 commodities correspond to more than one NAICS industry. For this purpose, agriculture production data from the U.S. Department of Agriculture Economic Research Service were used to determine a weight for each agricultural industry sector. Table A.2 presents the crop and livestock production value added to the economy in 2007 taken from the U.S. and State Farm Income Data Files, and the NAICS codes that correspond to each commodity. The values in Table A.2 were distributed between the NAICS codes so that a ratio could be computed for each NAICS with respect to the SCTG2 commodities. In Table A.3, fruits and tree nuts corresponds to two industry sectors, NAICS 1113A0 and 111335 (fruit farming and tree nut farming respectively), and with further data from the Farm Income Data Files, it was determined that fruit farming corresponds to 77 percent of the value and tree nut farming to 23 percent. In the remaining cases, the production values were divided equally between the industries. Table A.3 displays the distributed values to each industry and the computed weights/ratios.

Table A.2 Value Added to Economy by Agricultural Sector

Commodity	Value 2007[a] (In billions of dollars)	NAICS
Value of crop production		
Food grains	13.4	1111B0
Feed crops	42.3	1111B0, 1111A0
Cotton	6.5	111920
Oil crops	24.6	1111A0
Tobacco	0.8	111910
Fruits and tree nuts	18.5	1113A0, 111335
Vegetables	19.3	111200
All other crops	25.3	111400, 1119A0, 1119B0
Value of livestock production		
Meat animals	65.1	1121A0, 112A00
Dairy products	35.5	112120
Poultry and eggs	33.1	112300
Miscellaneous livestock	4.9	112A00

[a] U.S. Department of Agriculture, Economic Research Service, "United States and State Farm Income Data Files," http://www.ers.usda.gov/Data/farmincome/finfidmu.htm. Internet release date: December 15, 2010.

Table A.3 Employee Output and Value Added to Economy by NAICS

NAICS	Industry Description	Employee Output	Value (in Billions)	Weight	SCTG2	Commodity
1111B0	Grain farming	$50,000	$34.55	1.00	02	Cereal grains
1111A0	Oilseed farming	$100,000	$45.80	0.39	03	Other agricultural products
111200	Vegetable and melon farming	$100,000	$19.30	0.17		
1113A0	Fruit farming	$100,000	$14.26	0.12		
111335	Tree nut farming	$50,000	$4.26	0.04		
111400	Greenhouse, nursery, and floriculture production	$100,000	$8.43	0.07		
111910	Tobacco farming	$100,000	$0.81	0.01		
111920	Cotton farming	$50,000	$6.47	0.06		
1119A0	Sugarcane and sugar beet farming	$50,000	$8.43	0.07		
1119B0	All other crop farming	$50,000	$8.43	0.07		
1121A0	Cattle ranching and farming	$100,000	$34.99	0.25	01	Live animals and live fish
112120	Dairy cattle and milk production	$100,000	$35.45	0.26		
112A00	Animal production, except cattle and poultry and eggs	$100,000	$34.99	0.25		
112300	Poultry and egg production	$100,000	$33.12	0.24		

Next, the agricultural sector outputs were converted into agricultural employment. The FAF3 commodity-based agricultural production data by region was converted to industry production by region using the crosswalk (SCTG to NAICS) and the industry weights. The result is the production value by agricultural sector and FAF zone. With an assumed agricultural output per employee, the agricultural employment per FAF zone is calculated.

Table A.3 shows the output per employee assumed for each industry. The research conducted showed that the average agricultural output per employee in the U.S. varies depending on the region. The U.S. Bureau of Labor Statistics states that the average agricultural output per employee is $43,177. In the State of Arizona for the three-year average 1999-2001 the agricultural employee output was $61,616. And, in Monterey County, California the output per employee reaches $183,331. For this estimation, the employee output was assumed to be $50,000 or $100,000 depending on the value of the industry.

A.1 Example for FAF3 Zone: "Remainder of California"

The methodology described above was applied to each FAF zone. The following example for the "Remainder of California" region illustrates the procedure. The first step was to extract the agricultural production values from the FAF3 dataset. Table A.4 shows the values for the agricultural commodities, i.e., live animals and live fish, cereal grains, and other agricultural products. A total of $33.9 billion worth of agricultural commodities were produced in this region in 2007.

Table A.4 Production Values by Commodity for "Remainder of California"

		Value (in Millions) by SCTG			
FAF3 Zone ID	**Description**	**01 – Live Animals and Live Fish**	**02 – Cereal grains**	**03 – Other agricultural products**	**Total**
069	Remainder of California	$3,840	$3,137	$26,962	$33,939

These values were then distributed by industry sector (NAICS) with the ratios calculated in Step 2 of the methodology. Table A.5 shows the production value and employment by agricultural industry for the "Remainder of California." It is estimated that in this region the agricultural industry employs 434,783 people.

Table A.5 Employment by Agricultural Industry for "Remainder of California"

SCTG2	NAICS	Industry Description	Employee Output	Ratio	Value (in Millions)	Employment
03	1111A0	Oilseed farming	$100,000	0.39	$10,627	106,273
02	1111B0	Grain farming	$50,000	1.00	$3,137	62,747
03	111200	Vegetable and melon farming	$100,000	0.17	$4,478	44,784
03	1113A0	Fruit farming	$100,000	0.12	$3,309	33,093
03	111335	Tree nut farming	$50,000	0.04	$988	19,770
03	111400	Greenhouse, nursery, and floriculture production	$100,000	0.07	$1,956	19,564
03	111910	Tobacco farming	$100,000	0.01	$189	1,887
03	111920	Cotton farming	$50,000	0.06	$1,501	30,012
03	1119A0	Sugarcane and sugar beet farming	$50,000	0.07	$1,956	39,128
03	1119B0	All other crop farming	$50,000	0.07	$1,956	39,128
01	1121A0	Cattle ranching and farming	$100,000	0.25	$970	9,698
01	112120	Dairy cattle and milk production	$100,000	0.26	$983	9,825
01	112A00	Animal production, except cattle and poultry and eggs	$100,000	0.25	$970	9,698
01	112300	Poultry and egg production	$100,000	0.24	$918	9,178
				Total "Remainder of California"	**$33,939**	**434,783**

Appendix B: Quick Reference