Modeling complex Megaregion systems: Horizontal and Vertical Integration for a Megaregion Model

Rolf Moeckel1, Sabyasachee Mishra2, Fred Ducca3 and Tara Weidner4

1National Center for Smart Growth, University of Maryland
2Department of Civil Engineering, University of Memphis
3National Center for Smart Growth, University of Maryland
4Oregon Department of Transportation (ODOT)

1rolf.moeckel@udo.edu, 2smishra3@memphis.edu, 3fducca@umd.edu, 4tara.j.weidner@odot.state.or.us

Abstract

Megaregions are considered to be the new geography that may become the nation’s operative regions, according to the March 2010 U.S. Federal Highway Administration (FHWA) Strategic Plan. Such regions need analysis tools to evaluate scenarios and their regional impacts. These tools have to cover areas larger than covered by the typical Metropolitan Planning Organization (MPO) or State Department of Transportation (DOT) models. In addition, such tools need to integrate systems beyond transportation, such as economic, land-use or environmental modules. This paper reviews a megaregion case study, funded by the Federal Highway Administration’s Exploratory Advanced Research Program, using an analysis framework to address the impact of high energy prices in the Chesapeake Bay Megaregion around Washington D.C., USA. The paper describes the model, focusing on horizontal and vertical integration issues of various modules. A high-energy price scenario is analyzed showing the need for action at the megaregional scale.

Keywords: Multi-scale modeling, megaregion modeling, transportation/land use integration

1. Introduction

Megaregions across the United States have been analyzed for decades. The Boston-Washington corridor, the Chicago Metropolitan Region and the Los Angeles Basin are prime examples of frequently studied megaregions [1]. In Europe, one early megaregional concept, the Blue Banana, covers an arc stretching from Manchester in Northern England to Milan in Northern Italy. The Blue Banana was later rejected as being too simplistic (as it also covered the highly rural Alps) and the European Grape was proposed as an alternative [2], where every single grape abstracted a different European megaregion.

Formal megaregional arrangements are rare, despite the fact that political borders do not confine regional interactions and associated impacts. One notable quasi-megaregional agreement in the U.S. is the I-95 Corridor coalition, a group of transportation agencies and toll authorities along the U.S. East Cost from Maine to Florida. Being a volunteer and consensus-driven organization, however, the coalition is limited to actions that support the interests of each player involved. Megaregional organizations with executive power do not exist in the U.S. to date.

On the other hand, Megaregions comprise the economic engine of the U.S., forecasted to contain half the nation's population growth and perhaps up to two-thirds of its economic growth by 2050 [3]. Supporting the megaregions’ economic competitiveness domestically and abroad is a key concern given increasing global competition and
international trade. A primary justification for addressing policy issues at a megaregional scale as opposed to the metropolitan scale is that regional economic activities are increasingly linked. Economic shocks to one metropolitan area result in spillovers, both positive and negative, to adjacent metropolitan areas. As a consequence, the resulting environmental and social impacts associated with such activities similarly spill across metropolitan areas. Furthermore, as pointed out by Christaller [4], Lösch [5], and Ross and Woo [6], individual cities are part of larger systems that are linked by inter-city trade hierarchies.

Policy issues exhibiting the following characteristics are most appropriate for a megaregional governance framework.

- Investments involving large-scale economies, which are exhausted at the megaregional scale.
- Issues for which public sector demand is relatively homogeneous at the scale of the megaregion.
- Issues that involve a redistribution of resources across metropolitan areas or states but which benefit from local (megaregional) knowledge regarding the nature of the redistribution.
- Issues that can be addressed with low administrative costs at the megaregional scale. If there are economies of scale for administration, megaregional governance would be preferred to local governance arrangements.

Megaregional frameworks are most relevant for issues involving large spillovers that extend beyond existing local, regional, and state governance arrangements but not to the scale of the entire nation. Even if the spatial extent is contained within one state, megaregional governance arrangements are beneficial to address the externalities.

Many planning issues in the domains of environment, transportation, society and economic development meet these criteria and deserve megaregional attention. There are three methods to predict megaregional development. Behavior can be observed (observed preference) or people can be asked (revealed preference), and this behavior can be extrapolated into the future, or a mathematical model can be implemented. Out of these three alternatives, a model is the only option that allows studying the megaregional development under different scenarios. Given the geographic extend of megaregions, observing or interviewing is particularly challenging. A model allows capturing the most relevant interactions to analyze megaregional development under different scenarios. In addition, a model can also capture interactions between the megaregion and national and world economies, which would be very difficult to capture through surveys.

This paper provides a review of the literature (Section 2) and introduces the Chesapeake Bay Megaregion (CBM) around Washington D.C. (Section 3). Following, the framework for a megaregional model is described (Section 4), paying special attention at the multi-dimensional integration of modules and geographies (Section 5). The CBM model is described as an implementation of this megaregion modeling framework. Simulation results of a high-energy price scenario are presented (Section 6), and finally, implications for megaregions are discussed (Section 7).

2. Literature Review

Megaregions are defined in multiple ways in the literature. In the early 1960s, Gottmann [7] wrote a seminal book of the Northeastern Seaboard reaching from Washington to New York City, the largest urban conglomeration in the world at that time. An update was published by Gottmann and Harper [8], in which the megaregion remained mostly defined by population densities. A more common approach, adopted by the U.S. Census Bureau and Bureau of Economic Analysis, is to define regions in terms of labor
market commuting sheds, where the majority of workers commute to locations within the region for employment purposes. This approach is consistent with Hoover and Giarrantani’s [9] conception of a “nodal” region and Fox and Kumar’s [10] “functional economic areas,” where regional activities are oriented towards an internal nodal commercial business district, and there is a presumption of dominance of the node over the surrounding peripheral area [11]. Richardson [12] extends this concept to allow for polycentric regions with several nodes and several peripheries, a concept that is embodied in the U.S. Census Bureau’s current definition of a Combined Statistical Area (CSA).

The definitions proposed by recent authors differ in terms of the units of analysis that make up the underlying regions and how they are connected. The Regional Plan Association [13] along with urban planning graduate students at the University of Pennsylvania identified ten megaregions in the U.S., defined as those regions that are interconnected along at least one of the following dimensions: (1) environmental systems and topography, (2) infrastructure systems, (3) economic linkages, (4) settlement patterns and land use, and (5) shared culture and history. Hagler [14] proposes quantitative criteria to establish these linkages. An index was created with points assigned to counties according to whether the county was part of a core-based statistical area, had a population density exceeding 200 persons per square mile, and had increases in population, employment, or population densities exceeding certain thresholds. Ross et al., [15] proposed the following procedure to identify megaregions: (1) identify the core areas, (2) identify the boundaries of the areas of influence, (3) apply local characteristics, and (4) finalize the boundaries.

While there is a substantial body of literature on analyzing megaregions, little research has been conducted to modeling megaregions. The few applications published tend to focus on one aspect only, rather than integrating transportation, land use, demographic, economic and environmental models. For example, the Netherlands National Model System (NMS), often referred to by its Dutch name Landelijk Model Systeem (LMS), models travel demand for the entire Netherlands [26, 27], but only focuses on transportation. Another pure transportation model at the megaregional scale has been described by Zhang et al. [29]. They developed an impressive microscopic assignment model for a megaregion, but stopped short of integrating this model with land-use, economic or environmental components. A pure land use model for the evolving megaregion Dar es Salaam, Tanzania was developed by Lindner and Hill [30], but this model lacked the integration with a transportation model. The SASI (Spatial and Socio-economic Impacts of Transport Investments and Transport System Improvements) model is a notable exception as it models interactions of the economy with the transportation system for the entire European Union [28]. However, infrastructure benefits are only represented through accessibilities, they do not capture actual congestion on the transportation network. This paper fills this gap by developing a model for a megaregion that fully integrates economic, land-use, transportation and environmental modules.

3. Chesapeake Bay Megaregion

The Chesapeake Bay Megaregion (CBM) spans the area from the Southern border of Pennsylvania through Maryland and eastern Virginia to Norfolk and Virginia Beach. It includes all or part of six states: Pennsylvania, Maryland, Virginia, Delaware, West Virginia and the District of Columbia. The megaregion includes major cities such as Wilmington, Baltimore, Washington, Richmond, and the Norfolk area plus smaller areas such as Salisbury MD and Fredericksburg, VA. The CBM is defined by its primary environmental resource, the Chesapeake Bay. This Bay is the largest brackish water estuary in the U.S. and an important ecosystem for the region. The megaregion also consists of an advanced system of rail, ports, and highways that link labor markets and facilitate commodity flows, as well as linked labor markets that depend heavily on the
transportation and government sectors. **Figure 1** provides a map of the CBM along with major surface transportation infrastructure.

![Figure 1. The Chesapeake Bay Megaregion](image)

### 3.1 Transportation

The Chesapeake Bay Megaregion is a major transportation hub, and the transportation system not only connects it to the rest of the country but also helps to unite the megaregion. Interstate highways I-70 and I-270 connect the region to the west while I-81 and I-95 connect it to the north and south. In addition, Washington D.C., located near the center of the region, is a major rail hub with the Norfolk Southern, Chessie System and the AMTRAK Northeast corridor all passing through Union Station. The Ports of Baltimore and Norfolk-Hampton Roads connect to the land transportation network and the megaregion to the rest of world.

### 3.2 Economy

Economic factors define the size and character of the Chesapeake Bay Megaregion. According to Lang and Nelson’s [16] definition of megapolitan areas, this area is defined by the anchor metropolitan areas of Baltimore, Washington, Richmond, and Norfolk. Ross [17] defines a similar geography using cluster analysis. She finds that interactions within the Buffalo-Boston-New York-Philadelphia region and the Washington DC-Virginia region are stronger than between these two regions. This smaller geography also makes policy implementation more feasible, given that collaborative policy solutions require cooperation among a smaller number of states and local governments.

Strong economic linkages are demonstrated using a database based on IMPLAN\(^1\) input-output trade relationships. **Figure 2** presents the dollar value of freight flows between subareas of the megaregion. As can be seen, freight flows among the major cities

---

\(^1\) The data base consists of the dollar value of economic flows, disaggregated by 440 NAICS codes, between every county within the Chesapeake Bay Megaregion. This figure was constructed by taking the total dollar value of flows among all economic sectors and separating flows related to freight movements.
link the megaregion together. Not only are there strong economic ties along the I-95 corridor from Wilmington through Richmond, there are also major connections to Norfolk as well as connections to areas on the Eastern Shore of the Chesapeake Bay and to Western Maryland.

Figure 2. Dollar Values of Freight Flows Between Megaregion Subareas in 2007

The Chesapeake Bay acts as a major economic engine for the megaregion. The Bay supports outdoor recreation including sailing, swimming, boating and tourism. It is also a major food source, with oysters and crabs from the Bay shipped nationwide. The Bay is home to two major ports, Norfolk-Hampton Roads and Baltimore. These ports are well positioned to grow in importance as the Panama Canal is widened and the east coast develops more direct shipping linkages with Asia.

3.3. Political Linkages

Political linkages also help to define the Chesapeake Bay Megaregion. The region is a key part of the I-95 corridor coalition, a group of states from Maine to Florida concerned with traffic on the Interstate Highway I-95. Furthermore, the Chesapeake Bay Commission composed of the states Pennsylvania, Maryland and Virginia is a body set up to manage water and land resources within the Chesapeake Bay watershed. All areas within the megaregion are particularly concerned with the overall health of the Bay.

3.4 Policy Issues

By 2030, the Chesapeake Bay Megaregion is projected to be the fastest growing among all other subregions within the Northeast Corridor with 40.2 percent [16], giving rise to a range of growth-related policy challenges including traffic congestion and environmental pollution. Within Maryland and the Northern Virginia, land use and Smart Growth policies are of concern. The widening of the Panama Canal stands to redirect a portion of international freight flows to the CBM’s major ports of Baltimore and Norfolk. As a result, watershed protection, megaregional growth management, congestion pricing, land
use and port expansions are all issues that will likely rise in importance over the next several decades.

4. Modules

Given the geographic extend of a megaregion, a single scale model is insufficient to capture relevant activities, travel behavior and their impacts. Instead, a two-layer approach was chosen that distinguishes a megaregional layer represented in more detail and a national layer capturing relevant activities and flows outside of the main study area. Given the interactions between different megaregions nationally, and to some respect even globally, the two-layer approach facilitates to represent the study area with sufficient detail yet acknowledges that megaregions cannot be treated as monolithic islands.

Figure 3 visualizes the organization of the megaregion framework implemented for the Chesapeake Bay area. A series of economic models predict the growth and decline of different economic sectors based on assumptions of changes in the global economy. The economic models cover both the megaregional layer and the national layer to account for a global economy that affects growth and decline in the megaregion. This economic forecast is used to constrain both the land-use model and the transport model. The land-use model simulates changes in population and employment, including demographic/firmographic changes and household/employment relocation.

![Figure 3. Concept of the Integrated Megaregion Model](image)

The locations of population and employment are used in the travel demand model to simulate both person travel and freight flows. While the travel model has more detail in the megaregional layer, the national layer is relevant to cover long-distance traffic flows that may go into or through the megaregion study area. Accessibilities are fed back into the land-use model to influence land-use changes in the following simulation period, allowing land use to evolve over time. Finally, both traffic flows and the output of the land-use model are used in a series of indicator models to analyze the model results. The following sections describe the different modules in more detail.
4.1. Economic Model

To address the larger economic interest of a potential megaregion authority, the analysis framework emphasizes links with the economy, both nationally and locally. The megaregion model chain is driven by a national economic forecast component that predicts economic activity for large regions covering North America. This is used to drive long-distance freight flows, as well as provide control totals for employment and population to be allocated within the study area by the land use model. The national Computable General Equilibrium (CGE) economic forecasting model built by the INFORUM group at the University of Maryland is applied [18, 19]. The model employs inter-industry-macroeconomic general equilibrium models to examine past employment trends and to forecast future employment across 65 sectors of the economy. The primary model, LIFT (Long-term Inter-industry Forecasting Tool), uses econometric equations to predict final demand and output at the national level, based on inter-industry input-output relationships and value-added behavioral equations. The second component STEMS (State Employment Modeling System) allocates the national forecast to states, considering the industry’s mix of basic and non-basic employment and personal income. Basic employment is driven by LIFT national industry trends, while non-basic employment is driven by state-specific personal income forecasts. The land use model further allocates these state control totals to model zones.

To account for local transport impacts on the regional economy beyond the national economic drivers that have little knowledge of local conditions, an economic post-processor was developed. The post-processor combines knowledge of county-to-county goods movement by industry and its sensitivity to travel impedance (generalized cost including time, distance, and travel cost per mile) to determine the impact on an industry based on the local transport influence on their supply chain relationships (up and downstream). To do so, the following steps were completed:

- An empirically based supply chain dataset at the county level with full input-output relationships provided data on county-to-county financial flows for 2007 by industry (IMPLAN data). These data were aggregated into industry groups and isolated to sectors that ship products likely to travel by highway (by examining the purchased transportation inputs by mode, and the nature of the products shipped).
- The financial flows in dollars were converted to highway tonnage flows using ancillary data and mode split relationships. This resulted in truck tonnage as a function of county-pair characteristics, while retaining the industry connection.
- The travel model provided estimates of base and future years generalized impedances between county pairs, while the national economic model and land use model provided employment estimates for each county by industry category for base and future years.
- Based on the change in activity and travel impedance, estimates of the change in county-to-county tonnage flows were developed by industry group.

The above provides an economic post processor that allows identifying that under a given scenario Industry X will trade Y% more/less tonnage in counties A, B and C due to change in transport cost. This provides an industry and spatially explicit indicator of the impact of the scenario on the economy going beyond the top down economic assumptions. Ideally these impacts would influence the size and location of the region’s economic activity (feedback to an earlier stage in the model chain), which may be considered in a future upgrade. However, by stopping at this point before full feedback,

---

2 Freight Mode Choice model HaulChoice, (c) ECONorthwest 2010 based on IMPLAN Software v.3 (c) MIG, Inc. 2011, and county-level data provided by MIG, Inc.
the indicated economic impacts assume the industry and its suppliers do not relocate or change technologies to accommodate the transport costs. As such, it provides information for the megaregion to proactively plan to avoid such adverse industry impacts. The megaregion governing authority learns, which links/ portions of the transportation network need to be strengthened in order to keep the megaregion economically viable. The result is, in effect, a worst case bookend because (a) the highway network representation is not made a function of the expected economic growth (or decline) of the region, and (b) the base-year focus of industrial activity is held fixed, albeit scaled by the national economic forecast model. Although footloose industries may in fact relocate as highway facilities congest over time, astute improvements to the network afford the opportunity to mitigate this effect. Thus, the findings can better inform long-term planning and economic development.

4.2. Land Use Model

The framework for the economic/land use model is presented in Figure 4. The three stages are: (a) national level, (b) regional level and (c) local level. The first component was already described in Section 4.1 and provides control totals for the latter two, which operate only within the megaregional model layer.

- Regional Model: The disaggregation process at the county level begins with the input of historic employment data from the U.S. Bureau of Labor Statistics’ (BLS) Quarterly Census of Employment and Wages [20]. Employment data at the county level are used to calculate the share of employment of each county in the baseline years 2000, 2005 and 2009. The ratio of employment for each county is then extrapolated for five-year periods beginning in 2010 and ending in 2030 using an exponential smoothing method. Employment data from 101 BLS sub-sectors for the same period are then reconciled with 65 STEMS industries disaggregated to match the BLS sub-sectors. This process simply employs a straight average of the two employment ratios. The purpose of this method is to allocate future STEMS projections based on the historic difference in BLS data and STEMS projections. Based on these previous ratios from 2000, 2005 and 2009, the projection for the ratios until 2030 is extrapolated using OLS regression. The final disaggregation process takes the aggregate STEMS total employment number for each five-year projection cycle and allocates it by county and sub-sector based on the ratios developed in the previous steps.

Figure 4. Multi-layer Land Use Model
• **Local Model:** The local model results in land use outputs at the megaregional modeling zones (MMZ\(^3\)) level. The initial allocations are made based on transportation costs and the basic employment distribution. At the local level a Lowry model based allocation \([21]\) is used to allocate households and employment by five income categories from counties to MMZs. The input data of the model consist of (1) Friction factors by household income group, (2) Household and employment control totals at county level, and (3) Basic employment at MMZ level. The output includes (1) Households by five income categories, (2) Employment by four categories, (3) Error check files and plots. The model follows the Lowry approach in five steps:

- Step 1: Determine the shortest path travel times and calibrate the friction factors for base year by household type and income category.
- Step 2: Determine the household by income category from the county control total and the proportions derived from the travel impedances of basic employment (industrial).
- Step 3: Determine a specific type of non-basic employment using the county control total and the proportions derived from travel impedances of households of income category.
- Step 4: Go to step-2 and obtain a new set of households, repeat step-3 to obtain new set of employment.
- Step 5: Repeat step-2 through step-4 till convergence is achieved.

The model steps are repeated every five years until the horizon year 2030 is reached. The location of basic employment provides inertia in the local model’s allocation over time. To develop a functional megaregion model, a Lowry model was employed to obtain land use input in a timely matter. Since the megaregion size is a combination of states, obtaining detailed data to develop a land use model is a challenging task. In contrast, a Lowry model required less details of input but provides reasonable land use results efficiently and within computationally acceptable runtimes. However, to increase scenario sensitivities, more advanced and theoretically robust land use models should be considered in the future.

**4.3 Travel Demand Model**

In line with the land use model, the travel demand model works at two geographic layers (a megaregional with 2,075 zones and the national layer with 126 zones). While the local travel demand model is closer to a traditional Metropolitan Planning Organization Model, the national models are built from nationwide surveys for both person travel and truck long-distance travel. Travel demand of the two layers is combined in the traffic assignment process that assigns short- and long-distance traffic flows simultaneously in a multi-class assignment.

The megaregional travel demand model is built as a five-step aggregate model that includes trip generation, destination choice, mode split, time-of-day split and assignment. The original modeling concept was borrowed from the Baltimore Metropolitan Council model \([22]\), refined and applied for short-distance trips of 50 miles or less to the entire CBM study area. Trip generation rates were derived from a 2007 Household Travel Survey for the Baltimore and Washington D.C. metropolitan areas, as well as travel times, mode split and time-of-day split to fully calibrate the model to the core area of the CBM

\(^3\) MMZs are the polygon structures used in the megaregional model similar to TAZs in metropolitan transportation planning. MMZs in the megaregional model are equivalent to TAZs in high-density areas, and TAZs are nested in MMZs in low-density areas.
study area. Household Travel Surveys tend to have a sample size that is too small to fully calibrate mode choice options. Two on-board surveys that were conducted for planned transit expansions in the Baltimore and Washington regions were used to enrich data to calibrate mode choice.

This model includes a three-step local truck model based on the Quick Response Freight Manual (QRFM) published by the Federal Highway Administration [23], which was applied for local trips under 50 miles within the CBM study area. The comparison against VMT estimates revealed that the QRFM method, which is based on a 1992 truck trip survey from Phoenix, generates too many truck trips for the Chesapeake Bay study area. Therefore, parameters were scaled down to resemble local VMT truck estimates and truck count data. This model distinguishes three truck types, namely light (such as pick-up trucks or vans), medium (or single-unit) and heavy (or multi-unit) trucks.

At the national layer, the National Estimate of Long-Distance Travel (NELDT) has been implemented to simulate long-distance person travel greater than 50 miles [24]. The long-distance element of the National Household Travel Survey (NHTS) 2002, the most recent survey that covered explicitly long-distance trips in the U.S., provided data for long-distance travel behavior. To expand the survey to cover all long-distance travel in the U.S., air travel data from the Bureau of Transportation Statistics, which covers ten percent of all ticketed air travel passengers, was used as a control total for long-distance trips by air. Comparing the BTS data with the NHTS air travel records revealed that every 25,319th air passenger was included in the NHTS. Assuming that the NHTS is representative across all travel modes, this expansion factor was used to scale up the NHTS to represent the entire long-distance travel in the U.S. Auto trips were added to the multi-class assignment in conjunction with local traffic. Given the limitations of the sample size of the NHTS, some scaling was necessary to match count data at external entry points to the CBM study area.

For freight long-distance travel, commodity flows from the Freight Analysis Framework FAF3 were used for truck trips of 50 miles or more. FAF3 data are provided by 123 domestic FAF zones. Though these zones reflect centers of economic activity, the resolution is too coarse to model truck trips on a network with much more detail. Therefore, flows between FAF zones were first disaggregated to flows between counties and then further disaggregated to flows between model zones. To disaggregate commodity flows, make/use coefficients and employment by type were used to allocate flows to the most likely producers and consumers of every commodity.

Commodity flows between zones are converted into trucks trips using average payload factors provided by FAF. These factors describe how many tons of a given commodity are carried by a truck on average. To account for empty truck trips, an empty-truck trips rate is added globally to all truck trips. The model distinguishes two truck types, single-unit and multi-unit trucks. The truck type depends on the traveled distance. While single-unit trucks cover most of the short distance truck trips, multi-unit trucks dominate long-distance travel.

4.4. Environmental Model

Multiple indicator models are included that cover the environment, fiscal, and social impacts. The indicator models are used to estimate specific impacts from various policies using outputs from the transportation, land-use and economic models. The results of the indicator models are typically not fed back to the other model components but may be used to identify additional scenarios to test, such as economic, land use, or transportation actions necessary to keep below targeted indicator values.

**Indicator Model for Gaseous Emissions.** This model captures estimates of air emissions using the EPA Motor Vehicle Emission Simulator (MOVES) model [25]. The MOVES model uses VMT and link-level volumes and speed data output by the travel
model to estimate GHG and other mobile emissions. The model thus responds to changes in travel demand, vehicle fuel efficiency, VMT and/or speeds, reporting regional quantities of various emissions with each model run.

**Indicator Model for Water Quality.** This model captures the impact of alternative policies on water quality. A nutrient-loading model covers the portion of the megaregion draining into the Chesapeake Bay (although not the entire watershed). It forecasts the annual loads of nitrogen, phosphorus and sediments on the watershed. The model uses detailed land cover changes from a parcel-based land use model to identify changes in nutrient runoff experienced in each watershed. The model contains detailed data on ground classification for urban and agricultural land sub-classified into specific land cover categories. The model responds to changes in land cover, and thus any economic, transport, or land use policy.

**Indicator Model for Infrastructure Costs.** This model estimates state and local governments’ costs to provide public infrastructure in support of new development (e.g., roads, sewer, water). Established relationships between the current development and the provision of infrastructure are applied to project future improvements needed to satisfy additional activity. Data includes residential development classified by housing type, existing water and road infrastructure, property value trends and tax rates, among others. An infrastructure cost model uses relationships between urban/rural development and the provisions of infrastructure to forecast needs. The fiscal indicator model has been developed to reflect conditions and costs in Maryland. The model responds to economic, land use or transportation policies which impact land use.

5. Module Integration

Section 4 described four different groups of models that cover economic, land-use, travel and environmental aspects of the megaregion. These different modules have been tightly integrated to work in sync and pass information from one module to the next. The economic model sets the stage for the entire framework by defining growth or decline of population and employment as well as the amount of travel that is required to support this economy without exceeding the budgets set by the economy. The land-use model allocates households and employment to zones, which then are fed into the travel demand model to generate traffic. Traffic flows are used in the environment modules to estimate greenhouse gas emissions and other environmental impacts. And these flows of information happen at two different geographic layers, the megaregional and the national layer.

Therefore, this integration has two dimensions as shown in Figure 5. First, models of otherwise comparable phenomena may work at different geographical levels, such as integrating a travel model of the CBM area with a long-distance travel model (geographical integration). At a minimum, the output of the two models need to be combined, and often output from a model at one geographic layer directly influences the model behavior at another geographic layer.
Secondly, modules with the same geography but different modeling tasks need to be integrated horizontally, such as a transportation model and a land use model covering the same study area (component integration). The two modules are likely to improve by exchanging information. Each level is discussed separately in the following sections.

5.1. Vertical Integration Across Geographies

The CBM model is built as a geographic multi-layer approach integrating a national with a megaregional model. This two-layer approach requires close integration of modules to exchange data while avoiding double counting of modeled aspects. Interfaces were needed that facilitate the integration even under extreme scenarios. This integration needs to be bidirectional, with megaregion models utilizing information from national models and megaregion models providing information to national models. Megaregion models by their scope are most valuable to assess strategic region-wide policy actions, and should not try to replicate the urban-level forecasts that commonly are better represented by urban models.

Building modules that work at different geographies allows simulating similar tasks (such as person trips) with different modeling approaches catered to each geographic level. Every module may be designed differently, and the spatial resolution of different modules may differ to fit each model's purpose. While a destination choice model works well to distribute person trips at the local level, this module becomes difficult to apply simultaneously with both short-distance and long-distance trips. Thus, the same task of a person trip may be simulated with different methods at the megaregional and the national level. The spatial resolution is likely to be finer at the local level and much coarser at the national level. For a trip that stays within the study area, the detailed locations of origin and destination are of interest. For a trip that leaves the study area to a destination hundreds of miles away, the precise location of that destination most likely is irrelevant. In the CBM model, Washington D.C. is subdivided at the local level into 85 zones (called MMZ or megaregional model zones). At the national level, the finest resolution used is counties, with Washington D.C. being one jurisdiction at the county level. While a geographic distinction in different model layers is less relevant for urban models, this distinction is helpful when modeling larger study areas, particularly for megaregions.

The predecessor of the CBM model, the Maryland Statewide Transportation Model (MSTM), is integrated with the two urban models for Baltimore and Washington. This integration works in two dimensions. On the one hand, aggregated results of MSTM are compared with MPO model results to ensure consistency across geographies. Compared are number of trips generated, average trip length, mode split, and VMT by county. If the two layers agree across these dimensions, confidence of both the statewide and the MPO models is raised. In the interesting case that the two layers do not agree, seeking to understand the reasons for different model results may lead to intriguing conclusions on

![Figure 5. Two Dimensions of Model Integration](image-url)
which modeling approach works best for certain tasks. For example, the MSTM did not agree with the MPO models in terms of trip generation. After some research the team found that these models were using different household travel surveys, and the impact of using different surveys could be traced down from trip generation to the assignment.

For the integration in the other direction, model volumes from the MSTM are fed into some of the MPO models as traffic at external stations. In contrast to simple traffic counts, which only provide the total number of vehicles entering the MPO are at a certain location in the base year, the MSTM volumes specify how many of these external trips are through trips (providing the entry and exit point to and from the MPO area), and how many trips are internal-to-external/external-to-internal trips. For scenarios that may affect the travel behavior of long-distance travel, such as widening the capital beltway by adding more lanes, the MSTM provides updated external volumes under a given scenario. Figure 6 shows this integration geographically for the CBM model. Further integration of the CBM model with MPO models in the region is anticipated.

Figure 6. Three Geographic Layers of the CBM Modeling Framework

If trips are simulated at several geographic layers, special attention has to be paid to minimize inconsistencies at the boundaries between the layers. If the megaregional model had very small zones and the national model had very large zones, unrealistic behavior may be generated at the border. While outside the megaregion the model may only generate trips between zonal centroids that are fairly far apart, the model finds centroids that are close together inside the megaregion. This may lead to different trip length frequency distributions caused by the different resolutions in the zone system inside and outside the megaregion. The CBM model overcomes this inconsistency by applying different models to the more detailed megaregional zone system and to the coarser zone system at the national level. This way, both models can be calibrated to their respective zonal resolution, creating a more consistent trip length frequency distribution.

5.2. Horizontal Integration Across Modules

Within a geographic layer, a series of model components need to be integrated horizontally, including an economic model, a land use model, a person-travel demand
model, a truck model, and several environmental impact models. Every model is likely to benefit from (if not require) the integration with some or all other models. Figure 7 shows graphically the integration designed for CBM.

The economic model provides population and employment for the land-use model based on exogenously given growth. At this point, the economic model provides these data at the state level, and the land-use model allocates population and employment to the zonal level within these statewide constraints. At a future point, the economic model may be redesigned to provide the socio-economic control totals for the entire study area of the CBM model, allowing the land-use model to distribute population and employment entirely based on utilities of different locations, unconstrained by artificial state borders. The economic model also sets the stage for the transportation model, as it defines growth in long-distance person travel, long-distance truck travel and auto-operating costs based on exogenously given gas prices. The transportation model returns accessibilities to both the economic model and the land-use model. Accessibilities are considered to be one variable in predicting economic growth as well as the attractiveness of locations for households and firms to locate.

To calculate environmental impacts, the transportation model provides traffic volumes by vehicle type, time of day and speed at the link level. The land-use model provides land cover to the environmental impacts module, as different land cover types have different impacts on run-off water, fixed-source emissions and infrastructure costs. Conceptually envisioned, though not implemented yet, the environmental model will provide feedback to the land-use model on environmental quality. Environmental quality is a relevant location factor, as households enjoy living close to the shore of the Bay, close to parks, and away from highways and other sources of noise or pollution.

5.3. Module Interaction

At the outset, every module is calibrated independently. For example, the travel model was calibrated to match observed traffic flows in the base year, and sensitivities for policies in future years were tested. Similarly, the economic model was calibrated to observed employment numbers in the base year, and sensitivities to various economic scenarios were tested. After each module has been calibrated individually, they are
connected in one modeling framework as shown in Figure 7. After integration, no further calibration was necessary as the output of each module closely resembled base year conditions. Naturally, this connection leads to a certain disequilibrium over time. Congestion may change, and households may experience longer commute times. Some households may decide to relocate the following simulation period. But not all households that could improve their housing utility move at once, as the transactions costs are too high. Total convergence is not intended in such integrated models. The land use model will adjust location choices of households and employment based on changing accessibilities, and therefore, the modeling system is leaning towards an equilibrium. But land use and transportation will never reach a full equilibrium. Rather, some households relocate in one year, accessibilities change, and further households decide to relocate the following year based on updated accessibilities. In other words, every module moves towards equilibrium, but a true equilibrium is never reached because other modules changed in the meantime.

This continuous disequilibrium is assumed to better represent real-world development. Households are reluctant to relocate if accessibilities change, because moving is expensive. The only module that is assumed to be in true equilibrium is the traffic assignment, which iterates several dozen times to find an equilibrium. Drivers are assumed to adjust their routing choice almost immediately to changing traffic conditions, leading to a near-perfect equilibrium of traffic flows on the “average day” that is modeled. All other modules move towards an equilibrium but by design have built in a certain time lag to reach equilibrium.

5.4. Tightness of Model Integration

There is a wide range of tightness when integrating models. Models may share the same modules that are fed with different data to work at different geographies. Or, models may share some data that are reconciled to ensure consistency across each model. There are also different levels of how closely different modules may be integrated technically (Figure 8) [31, 32]. The most common is Integration Level 1. Every model runs independently. After a model has started, it reads the output data of several other models, completes its own calculations, writes new output data, and is closed. After one model has finished another model can start. Building one single piece of modular software (Integration Level 2) that contains all modules may be advantageous. Having all modules in one piece of software saves runtime because a large amount of data can be kept in working memory, saving read and write time. The integration into one piece of software that is likely to improve the runtime requires, on the other hand, a very close interaction between developers of all modules. The third level of integration runs all modules simultaneously. Events of each module are executed in random order, such as a person makes a trip to work, another household moves, a truck delivers groceries, a child is born, a person goes to the cinema, etc. This very close integration resembles how events happen in reality. So far, however, this level of integration rarely has been achieved in operational models.
The CBM model uses a combination of level 1 and level 2 integrations. While the economic and the land use model are stand-alone modules that are run consecutively, the long-distance person model, the long-distance truck model and the short-distance truck model are built as one single module that runs efficiently without time-consuming reading and writing of intermediate results.

6. Application

The CBM model was used to analyze the impacts of three scenarios: a base scenario for the year 2007 for validation purposes, a 2030 business-as-usual scenario (BAU), and a 2030 high-energy-price scenario (HEP). The single difference between the BAU scenario and the HEP scenario is the assumed price for gasoline. In the BAU scenario, a gas price of $2.90/gal is assumed, resulting in auto-operating costs of 12 cents per mile. The HEP scenario analyzes the impact of a gas price of $14.00/gal, which results in auto-operating costs of 42.2 cents per mile. The price of gas was set at a rather high level to test the resilience of the megaregion’s transportation and land use system response under an extreme scenario. The value of 42.2 cents/mile (in 2000$) was set at just over 3 times the INFORUM high-energy scenario fuel price and assumed a “shock” in prices, such that the population had only limited time to adjust the vehicle fleet. Still, average fuel efficiency in miles per gallon (mpg) was assumed to increase from 24.4 mpg in 2010 to 41.8 mpg in 2030.

The HEP scenario assumes energy prices rise suddenly to this higher price due to depletion of resources or market disruptions. No changes in land use are assumed, as the sudden price increase has not yet triggered long-term location decisions. The comparisons of these scenarios are presented for number of trips generated, the average trip length, the modes chosen and vehicle-miles traveled. This scenario section intends to show that the model provides reasonable responses to policy analyses. Since no considerable changes in land-use or economic data were observed in this sudden increase in energy costs, the analysis of this scenario focuses on impacts on the transportation system.

A Delphi Panel of experts in travel demand modeling was interviewed to develop elasticities in trip generation. The resulting sensitivities for trip generation range from -0.4% per 10% increase in gasoline prices for home-based work trips (least sensitivity) to -1.5% for home-based other trips (largest sensitivity). Additionally, long-distance person travel was assumed to be limited to a constant travel budget. Thus, increased travel costs led to a smaller number of trips and shorter distances for long-distance person travel.

Figure 9 compares the travel demand for the three scenarios by several trip purposes. Collectively, the number of trips in the 2030 BAU scenario is larger than the number of trips in the 2007 base year, largely due to population growth. Home-based work and non-home-based work trips show the least sensitivity to increased auto-operating costs under the 2030 HEP scenario. Home-based other trips as well as long-distance trips show a
larger sensitivity, as many of these trips are discretionary trips. Long-distance truck trips were assumed to have almost no sensitivity to high gasoline prices, as options to reduce trips are limited and a large share of increased transportation costs are passed on to the consumer.

Figure 9. Trip Generation by Trip Purpose

Figure 10 compares the average trip length in miles for the three scenarios analyzed. Trip lengths increased slightly from 2007 to 2030 in the BAU scenario, which is mostly due to employment and population growth resulting in job/housing imbalances throughout the region. The 2030 HEP scenario shows trip lengths that are notable smaller. Auto-operating costs are part of the generalized-cost function used to simulate destination choice. With higher gas prices, travelers seek destinations that are closer, and therefore, reduce the expenses for traveling.

Figure 10. Average Trip Length by Purpose
To understand the impact on mode choice, the transit share in the three scenarios is shown in Figure 11. As expected, home-based work and non-home-based work have the highest transit shares, mainly due to the transit-friendly location of many workplaces in downtown Baltimore and downtown Washington D.C. The BAU 2030 scenario shows a slightly lower transit share, as the region has grown in less transit accessible areas. In the HEP 2030 scenario, the transit share has approximately doubled in comparison to the BAU 2030 scenario, as higher gasoline prices make it more attractive to switch to transit where it is available.

Figure 11. Transit Share by Purpose

Vehicle-miles traveled (VMT) are a common way to assess the total travel on the highway network. Figure 12 compares VMT in the three scenarios by State, including Washington D.C. Note that these VMT numbers only include travel on the network covered by the megaregional model. Though the national layer of the model generates long-distance trips with trip ends outside of the megaregion, no short-distance trips are generated outside of the megaregion. In other words, total VMT in Pennsylvania is larger than shown in Figure 12, as only the part of Pennsylvania covered by the CBM study area is shown.

Between 2007 and 2030, VMT within the CBM region grew by 41%. This is due to a combination of growth in less transit-friendly areas and a general increase in population and income. The HEP scenario generates 19% less VMT than the BAU scenario. While higher gas prices will be a serious burden on many households and businesses, a positive side effect would be a relief in congestion and reduced transportation-related emissions.
Figure 12. Vehicle-miles Traveled by State, Including Washington D.C.

The scenarios show that transportation is significantly affected by changes in auto-operating cost. Effects include reduction in number of trips, shorter trip lengths and, where available, modal shifts from auto to transit.

7. Conclusions

The scenario results could only be achieved by integrating modules (horizontal integration) and geographies (vertical integration) into one megaregional model. Given the size of a megaregion, it is relevant to distinguish short-distance and long-distance travel. Even though every model gains from horizontal and vertical integration, a megaregional model benefits particularly due to the number of interactions and the geographic size. Integrating modules allows simulating complex interactions as found in the real world.

The model further allows analyzing the implications of scenarios on sub-regions and urban areas. By isolating those communities, population segments (e.g., by income, transit-accessible) and industries vulnerable to higher energy prices, the megaregion can use this analysis to identify policies to shore up these areas. This may help planning for a megaregion that is most resilient for economically competing under a high-energy price future.

The scenarios show a significant reduction in number of trips, trip length, auto share and vehicle miles traveled under the high-energy scenario. This would result in less congestion, fewer vehicle-hours of delay and a reduction of transportation-related emissions. However, such high surge in transportation costs will be a significant burden for many businesses and households, particularly those with lower incomes. In addition, air travel, which has not been modeled in this context, is expected to be affected severely by such high prices for gasoline. Though the exact impact is beyond this research, parts of the economy are likely to endure serious pressure under such a scenario as well.

Each module has been calibrated to match base year data for 2007, which included traffic counts for the transportation model, zonal households and employment for the land use model, and state employment and commodity flows for the economic model. The validation shows that the model is able to replicate base year conditions, even though the validation does not necessarily proof reasonable model sensitivities. Each module went
through a series of sensitivity tests to assess reasonable responsiveness heuristically. A more thorough analysis, however, would have been a back-casting study. In such analyses, the base year is set to a year in the past, such as 2000, and model results are compared to the most recent available current data, such as 2012. If the model back-casts changes from 2000 to 2012 reasonably, it would provide additional confidence that the model is capable of generating reasonable forecasts as well. Unfortunately, moving a model year is difficult as it requires collecting socio-economic data at the zonal level, updating highway networks, collecting traffic performance data (vehicle counts and VMT estimates), and gathering transit schedules and routings for a new model year. Given the size of the study area, data collection requires merging and reconciling datasets from various states, making it even more challenging to move the model year. As this research focused on the integration of various models, more effort was spent on developing methods of model integration, rather than spending too much time on data collection, back-casting and model validation.

Nevertheless, given the limited availability of natural resources and the global increase of demand for oil, a reduction in oil supply is inevitable in the future. Political unrest in some oil-producing regions further contributes to the increase of oil prices. The point of time by which the price for gasoline will rise to levels simulated in the HEP scenario is unknown. Regardless of whether it takes another ten years or fifty years, it is almost certain that the price for gasoline will rise to this level and beyond given the limitedness of resources and a worldwide growing demand. Though the impact will be severe both on the economy and on travel behavior, preparing a region for a high-energy price future may help mitigating negative impacts. Regions that are prepared are likely to cope more easily with the impacts of high gas prices. Changing land-use policies or redesigning the transportation system takes decades to show effects. The comparatively small price spike in 2008 happened within a matter of a few months, moving at a much faster speed than changes to land-use or to the transportation system may be implemented. Preparing a region before a steep price increase hits gives a region a competitive advantage in a high-energy price future.

Though some policies, such as environmental regulation or taxation, are most effective at the national or even global level, the megaregion is the appropriate entity to address many policies helping to prepare a region for increasing gas prices. If a single city limited urban sprawl, some households would move to a neighboring city without regulations, as housing prices might be lower. In contrast, if a megaregion defines land-use policies, the entire region will develop consistently. To a large extent the transportation system serves to connect urban regions. The megaregion is the most efficient level to implement region-wide transportation policies and to develop an integrated transportation infrastructure system that prepares the region for future challenges.

References

Authors
Rolf Moeckel, Dr.-Ing., works as a Faculty Research Assistant at the National Center for Smart Growth Research and Education at the University of Maryland. His main interest is to understand how computer modeling may support urban and regional planning. He has worked extensively on integrating land-use with transportation modeling. He has developed and implemented three land-use models
and continues researching how to integrate such models with travel demand models. To support regional and statewide planning, statewide or even megaregional travel demand models help decision makers prioritizing projects and policies. Dr. Moeckel has worked on nine statewide models across the U.S. Within these models, he has designed and implemented long-distance travel demand models for both person travel and freight travel. Prior to joining the National Center for Smart Growth, he worked with the Parsons Brinckerhoff as a Principal Professional Associate where he developed a wide range of models, including activity-based, mode choice, land use and freight models. He is a member of the TRB committee Travel Forecasting Resource (ADB45) and holds a doctorate in Spatial Planning from the University of Dortmund in Germany.

Sabyasachee (Sabya) Mishra, Ph.D., is an Assistant Professor with the Department of Civil Engineering at the University of Memphis. He is also affiliated faculty of the Intermodal Freight Transportation Institute at University of Memphis. His areas of expertise include Travel Demand Modeling, Transportation Planning and Policy, Econometric Modeling, Transportation Economics and Finance. Over the last few years, he has been involved in the development of statewide transportation models for various states as these models cater to short term and long term decision making of multimodal transportation system. He is a member of Transportation Economics (ABE20), and Freight Transportation Economics and Regulations (AT010) committees of the Transportation Research Board (TRB).

Fred Ducca, Ph.D., is Director of the Transportation Policy Research Group for the National Center for Smart Growth, research and Education at the University of Maryland. He has 27 years of experience in research, development, deployment and training in travel forecasting methods and procedures, including extensive experience providing technical assistance to state and local governments on travel forecasting issues. While with the Federal Highway Administration he managed the Travel Model Improvement Program (TMIP). This program advanced both the state of the art and state of the practice in travel forecasting. Dr. Ducca has a B.S. in mathematics from St. Peter’s College, an MBA from the University of Pennsylvania and a Ph.D. in City Planning, also from the University of Pennsylvania.

Tara Weidner, PE, is an Integrated Transportation Analysis Engineer in ODOT’s Transportation Planning Analysis Unit (TPAU). She has over 20 years of experience in modeling and analysis of multi-modal transportation systems. Her recent work has focused on quantifying the complex interactions of transportation, land use, and the environment. She joined TPAU about a year ago after working as a Senior Planner for Parsons Brinckerhoff (PB), where she led the consultant team partnering with the University of Maryland that worked on the FHWA Mega-regions project. This project included a case study using an expanded Maryland Statewide Model to look at the effect of high gas prices on the resilience of the Chesapeake Bay Mega-Region. Tara has played a key role in the integrated land use-transport models for Maryland as well as for the states of Oregon, Washington, and Ohio.