Mega-Regions develop as complex systems
Horizontal and vertical integration for a Mega-Region Simulation Model

Rolf Moeckel (moeckel@pbworld.com)
Parsons Brinckerhoff, 6100 Uptown Blvd. NE, Suite 700, Albuquerque, NM,
Phone (505) 878-6553, Fax (505) 881-7602

Sabyasachee Mishra (mishra@umd.edu)
National Center for Smart Growth, Preinkert Field House – 054, University of Maryland
College Park, MD 20742, Phone (301) 405-1945, Fax (301) 314-5639

Fred Ducca (fducca@umd.edu)
National Center for Smart Growth, Preinkert Field House – 054, University of Maryland
College Park, MD 20742, Phone (301) 405-1945, Fax (301) 314-5639

Tara Weidner (weidner@pbworld.com)
Parsons Brinckerhoff, 400 SW 6th Ave, Suite 802, Portland, OR 97204,
Phone (503) 478-2342, Fax (503) 274-1412

Abstract: Mega-regions are considered to be the new geography that may become the nation’s operative regions, according to the March 2010 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 mega-region 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 Mega-Region around Washington D.C. The paper describes the model, focusing on horizontal and vertical integration issues of various modules. A high energy price scenario is analyzed showing need for action at the mega-regional scale.

Paper length: 7,568 words and 12 figures

Acknowledgements: The work represented in this paper was funded through the “Mega-Region Travel Modeling Framework” project, sponsored by the Federal Highway Administration, Exploratory Advanced Research Program.
1. Introduction

Mega-regions 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 mega-regions (Florida et al. 2008). In Europe, one early mega-regional 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 covered most of the highly rural Alps) and the European Grape was proposed as an alternative (Kunzmann 2001), where every single grape abstracted a different European mega-region.

Formal mega-regional arrangements are rare in the U.S., even though regional interactions and associated impacts are not confined by traditional political borders. Indeed, the economic and environmental integration of metropolitan areas points to a need for new forms of governance to address issues that are mega-regional in scope. This raises the question, “Which policy issues are most appropriate for a mega-regional policy framework, and which are best left to existing local, state and federal institutions? And what analysis tools are needed?”

Mega-regions comprise the economic engine of the US, forecasted to contain half the nation's population growth and perhaps up to two-thirds of its economic growth by 2050 (Amekudzi, et al. 2007). Supporting the economic competitiveness of these regions domestically and abroad is a key concern given increasing global competition and international trade. A primary justification for addressing policy issues at a mega-regional scale as opposed to the metropolitan scale is that regional economic activities are increasingly linked in such a way that economic shocks to a given metropolitan area result in spillovers, both positive and negative, to adjacent metropolitan areas. As a consequence, the resultant environmental and social impacts associated with such activities likewise spill across metropolitan areas. Furthermore, as pointed out by Christaller (1933), Lösch (1954), and Ross and Woo (2009), individual cities are part of larger systems that are linked by inter-city trade hierarchies. Some policy issues exhibiting the following characteristics are most appropriate for a mega-regional governance framework:

- Issues involving large spillovers which extend beyond existing local, regional, and possibly state governance arrangements but not to the scale of the entire nation. In some cases, such issues may be appropriate for existing state-level governance arrangements. If a state is larger than the relevant mega-region, as in the case of Texas Triangle, mega-regional governance arrangements are appropriate when the spatial extent of the spillover is contained within the state, and there are benefits from relying on local (mega-regional) knowledge to address the externality.
- Investments involving large scale economies which are exhausted at the mega-regional scale.
- Issues for which public sector demand is relatively homogeneous at the scale of the mega-region.
- Issues which involve a redistribution of resources across metropolitan areas or states but which benefit from local (mega-regional) knowledge regarding the nature of the redistribution.
- Issues which can be addressed with low administrative costs at the mega-regional scale. If there are economies of scale in administration, then mega-regional governance would be preferred to local governance arrangements.
Several issues in the realm of environment, transportation, and economic development meet many of these criteria. There are three possibilities to predict mega-regional development. Behavior can be observed (observed preference), people can be asked (revealed preference), or a simulation model can be implemented. Out of these three alternatives, a model is the only option that allows studying the mega-regional development under different scenarios. Given the geographic extent of mega-regions, observing or interviewing is particularly challenging. A simulation model allows capturing the most relevant interactions to analyze mega-regional development under different scenarios. In addition, a simulation model can also capture interactions between the mega-region and national and world economies, something which would be very difficult to capture through surveys.

This paper provides a review of the literature and introduces the Chesapeake Bay Mega-Region (CBM) around Washington D.C. The CBM simulation model is described, paying special attention at the multi-dimensional integration of modules and geographies. Simulation results of a high energy price scenario are presented and implications for mega-regions are discussed.

2. Literature Review

Mega-regions are defined in multiple ways in the literature. 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 (1995) conception of a “nodal” region and Fox and Kumar’s (1994) “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 (Dawkins 2003). Richardson (1978) 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 (2006) along with urban planning graduate students at the University of Pennsylvania identify ten mega-regions in the U.S., defined as those regions which 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 (2009) 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, and population densities exceeding certain thresholds. Ross et al. (2008) proposes the following procedure for identifying mega-regions: (1) identify the core areas, (2) identify the boundaries of the areas of influence, (3) apply local characteristics, and (4) finalize the boundaries.
3. Chesapeake Bay Mega-Region

The Chesapeake Bay Mega-Region (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, West Virginia, Delaware, West Virginia and the District of Columbia. The mega-region includes major cities such as Wilmington, Baltimore, Washington, Richmond, and the Norfolk area plus smaller areas such as Salisbury Maryland and Fredericksburg, VA. The CMR is defined by its primary environmental resource, the Chesapeake Bay, an advanced system of rail, ports, and highways that link labor markets and facilitate commodity flows, and linked labor markets that depend heavily on the transportation and government sectors. Figure 1 provides a map of the CMR along with major surface transportation infrastructure.

3.1. Transportation

The Chesapeake Bay Mega-Region 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 mega-region. 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., 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 mega-region to the rest of world.

Figure 1: Chesapeake Bay Mega-Region
3.2. Economy

Economic factors define the size and character of the Chesapeake Bay Mega-Region. According to Lang and Nelson (2007) in their definition of megapolitan areas, one area is defined by the anchor metropolitan areas of Baltimore, Washington, Richmond, and Norfolk. Ross (2009) 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 data base based on IMPLAN\(^1\) input-output trade relationships. Figure 2 presents the dollar value of freight flows between subareas of the mega-region. As can be seen, freight flows among the major cities link the mega-region 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 mega-region sub areas (2007)](image)

The Chesapeake Bay acts as a major economic engine for the mega-region. 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 also 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.

---

\(^1\) The data base consists of the dollar value of economic flows, disaggregated by 440 NAICS codes, between every county within the Chesapeake Bay Mega-Region. This figure was constructed by taking the total dollar value of flows among all economic sectors and separating flows related to freight movements.
3.3. Political Linkages

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

3.4. Policy Issues

By 2030, the Chesapeake Bay Mega-Region is projected to be the fastest growing among all other sub-regions within the Northeast Corridor at 40.2 percent (Lang and Nelson 2007), 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 CMR’s major ports of Baltimore and Norfolk. As a result, watershed protection, mega-regional growth management, congestion pricing, land use and port expansion are all issues that will likely rise in importance over the next several decades.

4. Simulation modules

Given the geographic extent of a mega-region, 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 mega-regional layer represented in a more detail and a national layer capturing relevant activities and flows outside of the main study area. Given the interactions between different mega-regions nationally, and in some respect even globally, the two layer approach facilitates representing the study area in sufficient detail yet acknowledging that mega-regions cannot be treated as monolithic islands.

Figure 3 visualizes the organization of the mega-region model developed for the Chesapeake Bay Area. A series of economic models predicts the growth and decline of different economic sectors based on assumptions of changes in the global economy. The economic models cover both the mega-regional layer and the national layer to account for a global economy that affects growth and decline in the mega-region. 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 mega-region model

The location 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 mega-regional layer, the national layer is relevant to cover long-distance traffic flows that may go into or through the mega-region 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 describes the different modules in more detail.

4.1. Economic Model

To address the larger economic interest of a mega-region authority, the analysis framework emphasizes links with the economy, both nationally and locally. The mega-region 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. A proprietary national Computable General Equilibrium (CGE) economic forecasting model built by the INFORUM group at the University of Maryland is applied. 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
counties and ultimately to model zones. A detailed description of LIFT and STEMS can be found in the literature (McCarthy, 1991; INFORUM, 2010).

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 also 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). This data was 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 relationships2. 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 year 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, by industry group, were developed.

The above provides an economic post processor which allows us to say that under a given scenario Industry X will trade Y% less tonnage in Counties A,B and C due to change in transport cost. (e.g., Frederick County Virginia will trade 10% less agriculture with Montgomery County Maryland due to transport costs.) 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 early stage in the model chain), which may be considered in a future upgrade. However, by stopping at this point before full feedback, 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 mega-region to pro-actively plan in order to avoid such adverse industry impacts, informing the mega-region governing authority which links/portions of the transportation network need to be strengthened in order to keep the mega-region 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 locus 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.

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.
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, (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 mega-regional model layer.

- **National Model:** The national econometric model consists of two sub-models: (1) Long term Inter-Industry Forecasting Tool (LIFT), a macroeconomic input-output model operating at the U.S. national economy level forecasts more than 800 macroeconomic variables; that are then fed into (2) State Employment Modeling System (STEMS) to calculate employment and earnings by industry for all 50 states and the District of Columbia (described in more detail in section 4.1).

- **Regional Model:** The disaggregation process at the county level begins with the input of historic employment data from the Bureau of Labor Statistics’ (BLS) Quarterly Census of Employment and Wages. 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 properly 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 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.

- **Local Model:** The local model results in land use outputs at the mega-regional 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 (Lowry 1964) 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.

### 4.3. Travel Demand Model

The travel demand model works at two geographic layers, the mega-regional and the national layer. 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.

The mega-regional 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 (BMC 2007), 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 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 region were used to enrich data to calibrate mode choice.

---

\(^3\) MMZs are the polygon structures used in the mega-regional model similar to TAZs in metropolitan transportation planning. MMZs in the mega-regional model are equivalent to TAZs in high-density areas, and TAZs are nested in MMZs in low density areas.
This model includes a three-step local truck model based on the Quick Response Freight Manual (QRFM) published by the Federal Highway Administration (Beagan, Fischer et al. 2007), 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. Hence, parameters were scaled down to resemble local VMT truck estimates and truck count data. This model distinguishes three truck types, namely light (i.e. pick-up trucks, 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 (Moeckel and Donnelly 2011). The long-distance element of the national household travel survey (NHTS) 2002, the last publicly available survey that covered explicitly long-distance trips in the U.S., was used to extract 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 traffic in the U.S. Auto trips were extracted and 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 well 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 certain commodity can be 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 is chosen based 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 sustainability triple bottom line of environment, fiscal, and social impacts of importance to the CBM. 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: Gaseous Emissions.** This model captures estimates of air emissions resulting from various policy changes using the EPA Motor Vehicle Emission Simulator (MOVES) model (EPA 2011). 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 respond to changes in travel demand, vehicle fuel efficiency, VMT and/or speeds, reporting regional quantities of various emissions with each model run.

**Indicator Model: Water Quality.** This model captures the impact of alternative policies on water quality. A nutrient loading model covers the portion of the mega-region 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: 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 current development and the provision of infrastructure are applied to project future improvements needed to satisfy additional activity; assumes different levels of service for urban and rural areas. Data includes residential development classified by housing type; existing water and road infrastructure and capacities. Property value trends, tax rates, etc. 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 mega-region. These different modules need 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 mega-regional and the national layer.

Thus, 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 needs to be combined, and often output from a model at one geographic layer directly influences the model behavior at another geographic layer.

![Figure 5: Two dimensions of model integration](image)

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 multi-layer approach integrating a national with a mega-regional model. This two layer approach requires close integration of modules to pass on data required by each model, avoid double-counting of aspects simulating, and develop smooth interfaces that facilitate integration even under extreme scenarios. This integration should be two-way, with mega-region models utilizing information from other models and mega-region models providing information to other models. Mega-region models by their scope are most valuable in assessing strategic region-wide policy actions, and should not try to replicate the urban-level forecasts best left to MPO models.

Building modules that work at different geographies allows simulating similar tasks (such as person trips) with different modeling approaches catered to each level. Each 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 with both short-distance and long-distance trips using the same calibration results. Thus, the same task of a person trip may be simulated with different methods at the local and the regional level. The spatial resolution may be finer at the local level and much coarser at the regional 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 a hundred miles away, the precise location of the destination most likely is irrelevant. In the CBM model, Washington D.C. is subdivided at the local level into 85 zones (called MMZ or mega-regional model zones). At the national level, the finest resolution used is counties. While a geographic distinction in different model layers most likely is less relevant for urban models, this distinction is helpful when modeling larger study areas, such as a mega-region.
Further integration of the CBM model with MPO models in the region is anticipated. The predecessor of the CBM model, the Maryland Statewide Model (MSTM), is integrated with the two MPO 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, it is important to understand the reasons for different model results. 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 all the way from trip generation to the assignment. For the integration in the other direction, model volumes from the MSTM are planned to be 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 area at a certain location, 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 travel behavior long-distance travel, such as widening the capital beltway by adding more lanes, the MPO regions may consider implementing this scenario in MSTM to provide updated external volumes under a given scenario.

While this integration has not yet been realized for the CBM model, it is planned to integrate CBM with the MPO models of Baltimore MD, Washington DC, Fredericksburg VA, Richmond VA and Norfolk VA. Figure 6 shows this integration geographically.
If trips are simulated at several geographic layers, special attention has to be given to minimize inconsistencies at the border between the layers. If the mega-regional model had very small zones, and the national model had very large zones as spatial representation, pathological behavior may be generated at the border. While outside the mega-region the model may only generate trips between zonal centroids that are fairly far apart, the model finds centroids that are close together inside the mega-region. This may lead to different trip length frequency distributions that are solely caused by the different resolutions in the zone system inside and outside the mega-region. The CBM model overcomes this inconsistency by applying different models to the more detailed mega-regional 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

At the same 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) an integration with some or all other models. Figure 7 shows graphically the integration designed for CBM.

![Figure 7: Horizontal integration of modules for CBM](image)

The economic model provides population and employment for the land-use model based exogenously given overall 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 are of the CBM model, allowing the land-use model to distribution population and employment entirely based on utilities of different locations, unconstrained by artificial state borders. The economic model also set the state 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. Though not implemented yet, it is anticipated that the environmental model will provide feedback to the land-use model on environmental quality in subsequent model versions. 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. 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). 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, does its own simulation, 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 run 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 applied 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, a 2030 business-as-usual scenario (BUA), and a 2030 high-energy-price scenario (HEP). The only difference between the BUA scenario and the HEP scenario is the assumed price for gasoline. In the BUA 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 into auto-operating costs of 42.2 cents per mile. 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 travelled.

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 reduced number of trips and shorter distances of 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 BUA 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 may be 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 length 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 significantly 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 available.

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 mega-regional model. Though the national layer of the model generates long-distance trips with trip ends outside of the mega-region, no short-distance trips are generated outside of the mega-region. In other words, total VMT in Pennsylvania is larger than shown in Figure 12, as only the part of Pennsylvania covered by the mega-region 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 will be a relief in congestion and reduced transportation-related emissions.
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, shifting from auto to transit.

7. Conclusions

The scenario results could only be achieved by integrating modules (horizontal integration) and geographies (vertical integration) in one mega-regional simulation model. Integrating the transportation model with the land-use model allows triggering interactions, such as households moving further out if increased fuel-efficiency reduces travel costs. Given the size of a mega-region, it is relevant to distinguish short-distance and long-distance travel. Even though every simulation model gains from horizontal and vertical integration, a mega-regional 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-region and urban area. By isolating those communities, populations (e.g., by income, transit-accessible) and industries vulnerable to high energy prices, the mega-region can use this analysis to identify policies to shore up these areas; with the goal of planning for a mega-region that is most resilient for competing economically 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 will result in less congestion, fewer vehicle-hours 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 unknown, parts of the economy are likely to endure serious pressure under such a scenario.

Given the limited availability of natural resources and the global increase of demand for oil, a reduction in oil supply is unavoidable 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 twenty years or twice as long, it is almost certain that the price for gasoline will rise 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 this 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 mega-region is the ideal entity to address many policies helping to prepare a region for increasing gas prices. If one city limits urban sprawl, some households will move to a neighboring city without regulations, as housing prices might be lower. In contrast, if a mega-region defines land-use policies, the entire region will develop consistently. To a large extend, the transportation system serves to connect urban regions. The mega-region 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.

8. References


Highway Administration and Center for Quality Growth and Regional Development: Washington, DC.
Moeckel, Rolf, Donnelly, Rick (2011). *Nationwide Estimate of Long-Distance Travel (NELDT) Generating External Trips for Local Travel Demand Models*. Annual Meeting of the Transportation Research Forum, Long Beach, CA.