International Electricity Market Module

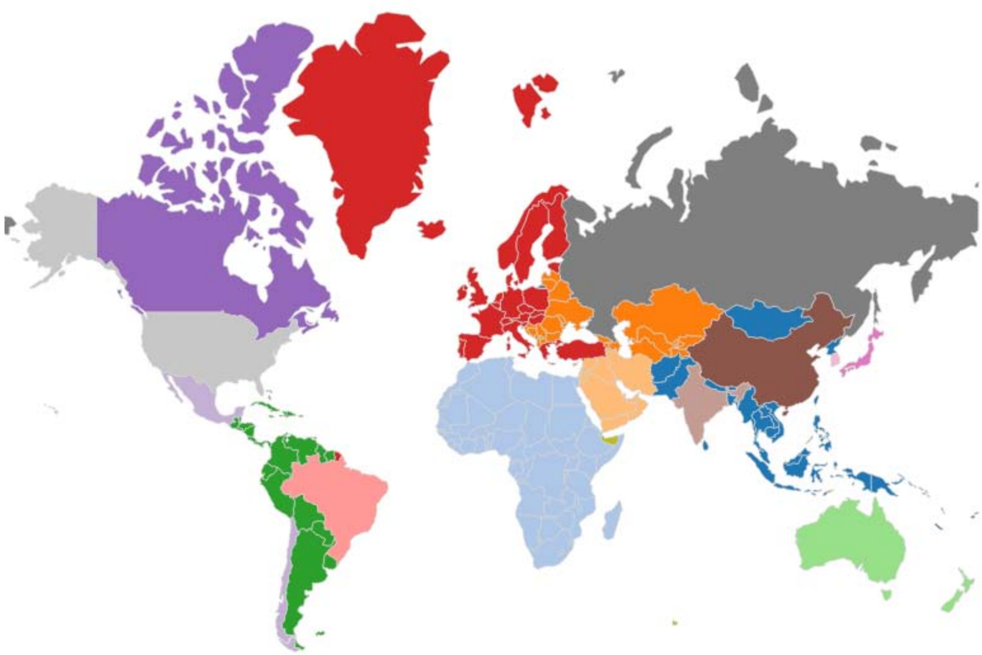
Component Design Report

OnLocation, Inc.

11/24/2015

Prepared For

Energy Information Administration



Deliverable No. CLIN No. 00001 - IEMM CDR

Page left intentionally blank

Contents

Executive Summary **1**

1 Overall Design **3**

1.1 Background 3
1.2 Flexible Design 4
1.2.1 Regional Disaggregation 4
1.2.2 Balanced Solution Techniques 5
1.2.3 Responsive Software Design 5
1.3 Comparisons to Other Modeling Approaches 5

2 Input Requirements **7**

2.1 Passed from Other Modules 7
2.1.1 Regional Sector Demands 7
2.1.2 Regional Fuel Prices 7
2.1.3 Macroeconomic Data 8
2.2 Data Coming from Exogenous Input Files 8
2.3 Evaluation of Existing Data Sources 9

3 Data Assessment and Sources **10**

3.1 Economic and Demographic Data 10
3.2 Investment Environment (Corporate, State Owned, Other) 11
3.3 Factors Influencing Load Shapes and Demand 12
3.3.1 Load Shape Archetypes 12
3.3.2 Differentiating Load Shape Archetypes and Load Duration Curves 13
3.3.3 Data Sources 14
3.4 Regional Policies 14
3.4.1 Corporate Income and Other Investment Related Taxes 14
3.4.2 Investment Incentives and Subsidies (positive or negative) 15
3.4.3 Taxes (VAT and other) 16
3.4.4 Discount Rates and Risk Premiums 16
3.5 Resources/Access to Fuels 17
3.5.1 Renewable Energy 17
3.5.2 Conventional Fuels (aka, coal, oil, natural gas, hydroelectric) 18
3.6 Technology Characterization and Regional Applicability 19
3.6.1 Initial Capacity and Characteristics 19
3.6.2 Cost and Performance Data Characterizing New Generation Technology 20

4 Methodology Description 22
4.1 Model Objective(s) 22
4.2 Regional Definition and Modeling Methodology Selection 22
4.3 Common Elements 25
4.3.1 Technology Representation 25
4.3.2 Renewable Resources 26
4.4 Optimization Model Structure 26
4.4.1 Load Representation 26
4.4.2 Basic Optimization Model: Capacity Planning and Dispatch 27
4.4.3 Alternative Combined LP Solution Approaches 36
4.5 Detailed Simulation Model Structure 37
4.5.1 Load Representation 38
4.5.2 Capacity Planning: New Capacity and Retirements 38
4.5.3 Technology Shares 39
4.5.4 Dispatch 44
4.6 Simple Simulation Model Structure 45
4.7 Estimating Sector Prices 45
4.8 Aggregating to WEPS+ Regions 48

5 Outputs 49
5.1 Passed back to other modules 49
5.2 Additional Electricity Reports 49

6 Running the Model 50
6.1 Standard WEPS Flow: Iteration to establish supply-demand balance 50
6.2 User Interface 50
6.2.1 Model Generation 51
6.2.2 Custom Reporting 52

7 Implementation Approach 53
7.1 Software Considerations 53
7.2 Object-Oriented Programming in Python 54
7.3 Preparing the Input Data 54
7.4 Auto-Generation of Model Objects 55
7.4.1 Simulation Objects 55
7.4.2 Optimization Objects 55
7.5 Aggregation and Reporting 56

8 Uncertainty and limitations 57

9 Recommendations..58

9.1 Pre-Development Software Design Review 59

9.2 Phase I Prototype 59

9.3 Phase II 60

9.4 Phase III 61

Appendix A: MODEL COMPARISONS 63

Appendix B: OPEN SOURCE PACKAGES FOR OPTIMIZATION PROBLEMS 67

Appendix C: OPEN-SOURCE SOLVERS 68

List of Tables

Table 1: Summary of Illustrative Initial Shares of Africa Demand 23

List of Figures

Figure ES-1: Block Diagram of Model Components 2
Figure 3-1: Alternative Residential Archetypes 13
Figure 4-1: User Interface to Establish Sub-regions 23
Figure 4-2: Details Behind Illustration of Resultant Breakup of Africa 24
Figure 4-3: Decomposition Illustration 37
Figure 4-4: Example with four load segments 38
Figure 4-6: Two technology logit example with lambda 40
Figure 4-5: Example of logit distribution 40
Figure 4-7: Flow of capacity expansion within a load segment 43
Figure 6-1: User Interface to Establish Sub-regions (repeated from above) 50
Figure 6-2: Contrasting Wealth Concentration with Electricity Use Per Capita 51
Figure 7-1: Block Diagram of Model Components 54
Figure 9-1: Overview of Responsive Design Process 59

Executive Summary

The Energy Information Administration (EIA) faces a number of challenges in developing its new International Electricity Market Model (IEMM). Among the more critical are:

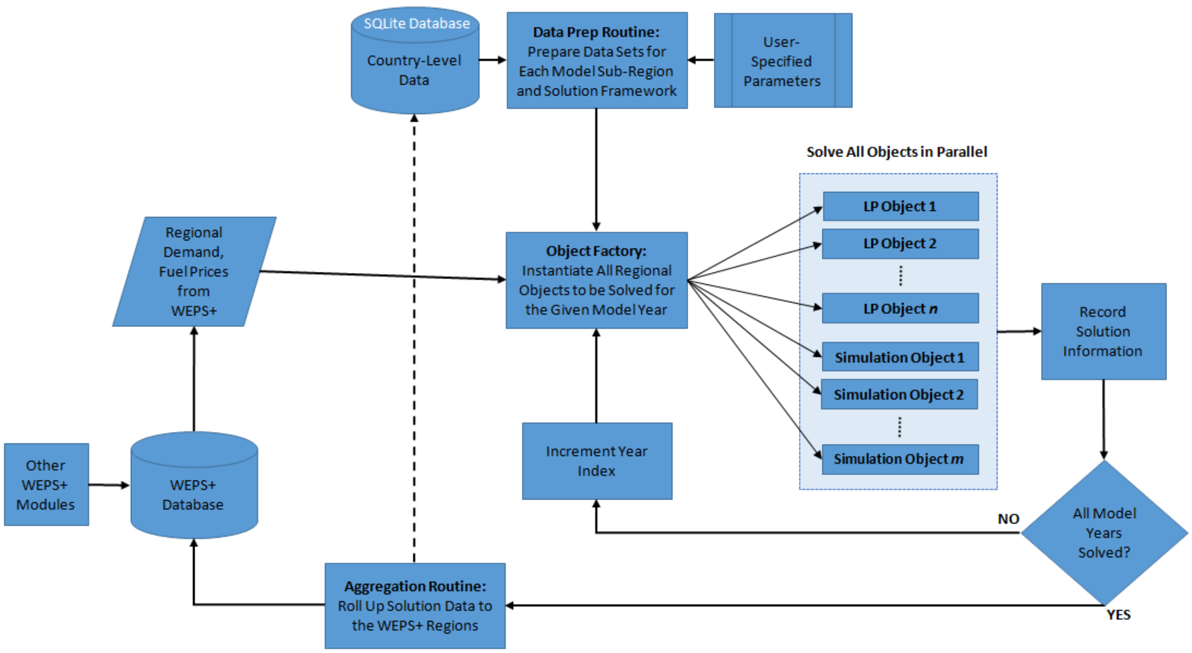
- Evolving policies designed to address climate change
- Evolving technology choices
- Evolving patterns of deployment of electricity-generating technologies and in particular distributed generation

Any new model that EIA develops should be able to address each of these in a flexible way. Equally as important, the model needs to be built on a platform that allows for rapid change to incorporate the latest technology and deployment approaches that arise over time. This Component Design Report (CDR) describes in detail an approach to the IEMM that will allow EIA to meet those challenges and to be readily adapted based on new information (technology, deployment strategies and policies).

Figure ES-1 provides an overview of the components of the design that will be described in detail in the CDR. The model design includes a front-end preprocessor that manages the data and is used in the regional definition formulation. The core of the model includes three different techniques for selecting capacity, fuel use, electricity prices and other key outputs of the IEMM. For those regions/sub-regions with large capacity existing in established power grids, we have designed a linear programming (LP) optimization approach. For regions/sub-regions that have significant but lesser needs for capacity and we have limited data describing the existing capacity, we have designed a detailed simulation model. Finally, for those areas/loads that have limited data and serve smaller, lower income sub-regions, we have a simple simulation model. The back-end processes data for integration with the rest of WEPS+.

The design of this modeling approach addresses all of EIA's needs. The approach is based on the costs and performance of competing technologies. It is designed to allow flexibility in addressing emerging distributed generation deployment concepts and accommodate regions and /or sub-regions with limited data. It is designed to be able to address policies such as climate change initiatives, renewable portfolio standards, and others.

Figure ES-1: Block Diagram of Model Components



The design of the modeling approach is flexible and responsive to change. The model is data driven. As new and more information becomes available in various regions of the world, the design allows for this information to be readily incorporated into the definition of the model objects. The design allows for multiple solution techniques, recognizing varying degrees of information and appropriateness to the sub-region defined. Additional solution techniques can easily be added or the existing modified specifications. By combining a data-driven process with a flexible solution choice, the design inherently allows the variety of regional circumstances to be captured in the electricity forecast.

The design of the modeling approach will allow the resultant model and database framework to be maintained, modified and extended by EIA, using software currently in use at EIA and/or publically available. The design uses SQLite (which is currently used by the WEPS+ team) to contain all of the data (raw data can still be stored in Excel or csv files). The main modeling routines will be written in Python which is the primary language used by EIA in WEPS+ and by other EIA analysts for data manipulation. The optimization solver recommended is open sourced, so EIA can have complete control over it. If, however, EIA prefers to use one of its currently licensed solvers, that can also be easily accommodated.

This design addresses all of EIA's needs, it is flexible and responsive to change and it can be maintained by EIA.

1 Overall Design

1.1 Background

A core question for the International Electricity Market Module (IEMM) is how will the global electricity supply and its related fuel consumption evolve over the next several decades as income rises in many regions and electricity demand grows? As articulated by the Energy Information Association (EIA), the most important alternative scenario for the modeling is the ability to represent responses to global energy policy initiatives and in particular those aimed at mitigating carbon emissions from the power sector. The proposed model structure is based on economic principles in projecting technology selection and dispatch, whether the simulation or optimization technique is used. As a result, policies such as carbon fees, fuel or technology subsidies can be directly represented. The technology choice can also be modified to represent renewable or other technology standards. These policies can be implemented at the sub-regional level or, if desired, globally, such as for a carbon fee. The questions that the model will be asked to address should shape the design and parameterization of any model.

The model will also have the technology detail to respond appropriately with lower carbon emitting options, such as various renewable types and nuclear power, taking into consideration their unique characteristics, such as intermittency in the case of some renewables, and political viability. In addition, economically motivated retirements will be built into the design. Because technology cost and performance assumptions can be critical to the outcome of climate policy scenarios, the model inputs will be constructed so that these are easy to modify in a comprehensive and consistent way for sensitivity cases.

Finally, a key element of modeling the electricity markets globally is to recognize that the traditional central power grid will need to evolve to include smaller grids and micro-grids. While certain regions like Europe, the US, Canada, and others will continue to build out around the central power concept, even these regions will feel the pressure based on reliability, security and local interests to allow for and expand the integration of micro-grids into their larger systems. Regions like Central and South American (CSA), the Other Asian States (OSA) and Africa will need a combination of approaches to meet the needs of their citizens going forward, particularly if an overarching carbon policy is to eventually be achieved.

1.2 Flexible Design

The modeling approach to the IEMM we are proposing is a data-driven concept built around flexibility, transparency and an responsive software development process. While there is a myriad of approaches to addressing this modeling problem, we believe the combination of a geographical information system, auto-generation of model objects, and customized solutions algorithms will provide:

- **Ease of Use and Maintenance:** By storing the key data in a GIS-focused data base, the datasets can be analyzed by modern geographic information systems that allows the user to inspect, verify and validate the datasets. Updating the inputs to the model then focuses on updating and maintaining the GIS data. Further, this approach lends itself to distributed upkeep by allowing regional/country experts to keep fresh each dataset, technology experts to maintain those data sets, etc. Finally, to the extent applicable, Application Program Interfaces (APIs) can be employed to periodically refresh the datasets. By using the spatial anchoring and GIS framework we can auto-generate the abstract logical model objects system which then can be "solved" by the IEMM solution algorithms.
- **Transparent and Easily Interpreted Object Model Formulation:** The output of the GIS framework is a set of model objects (classified by solution technique) matched to the user's criteria addressing size, scope, economic, technology characteristics and other attributes. The resultant set of objects can be displayed and investigated by the modelers to validate and verify the model specification.
- **Flexible Model Framework:** By using a set of solution algorithms matched to the data we can ensure consistent, logical solutions. Further, this object-oriented modeling framework allows us to expand the number of modeling objects as dictated by the user and the analysis being performed. Finally, the solution algorithms can be easily enhanced or expanded due to the modular (object) way of implementing the modeling framework without disturbing the modeling framework, allowing for parallel testing of new features.

Taking this approach the model will support evolutionary development, early prototype delivery, continuous improvement, and encourage rapid and flexible response to change.

1.2.1 Regional Disaggregation

As described more fully in Section 4.2 below, our approach is centered on the flexible definition of model objects that decompose the WEPS+ aggregate regions into logical abstract objects representing shares of WEPS+ regional kWh demands to be passed to the selected solvers. The fractional shares will be based on user-defined country level attributes (e.g. size of grid (GW), income levels (GDP per capita), use per customer (kWh per capita), GINI to allocate income/kWh to urban areas, technology and fuel availability, etc.). In its most simple form, the regions can be left aggregated and sent to any of the solvers. A more likely case is where from one to ten potential differentiated shares would be sent to one of the three currently anticipated solution approaches.

1.2.2 Balanced Solution Techniques

Section 4 provides for three different techniques for selecting capacity, fuel use, electricity prices and other key outputs of the IEMM. For those regions/sub-regions with large capacity existing in established power grids, we have designed a linear programming (LP) optimization approach. For regions/sub-regions that have significant but lesser needs for capacity and we have limited data describing the existing capacity, we have designed a detailed simulation model. Finally, for those areas/loads that have limited data and serve smaller, lower income sub-regions, we have a simple simulation model. The model front-end software will allow the user to easily test alternative combinations of sub-regions/solvers based on a combination of market (size and income levels), technology and fuel availability.

1.2.3 Responsive Software Design

Our approach is to use off the shelf core software (e.g., Python and associated libraries), in conjunction with GIS software and SQLite database software to build out the IEMM (see section 6.2 below, for more detailed discussion.) We will have separated the raw data from the model inputs. The solution approaches will stand by themselves and be testable outside the full implementation. The IEMM specific report writing and output databases will be separated into its own module. By taking this approach, we will be able adapt and evolve the design as new data and concepts for solutions become available. For those regions with extensive data available we will be able to take advantage of that data and use it to build out the appropriate LP optimization framework. Further, due to the flexible object oriented framework, we will be able to test alternative levels of aggregation representing alternative sub-grids, etc. As the solution algorithms will be contained methods, new ones can easily be added to the framework. This approach will allow us to provide early delivery of core pieces for testing and evaluation by EIA, continuous improvement, and it also encourages rapid and flexible response to change both during the development cycle and long-term.

1.3 Comparisons to Other Modeling Approaches

As EIA embarks on constructing a new international electricity model for WEPS+ it is useful to examine other global electricity models that are similar in the types of forecasts and policy questions they address.

A short description is provided in Appendix A of five global integrated energy models along with their key attributes.[1] Four of the five models are simulation models, and the remaining model (MARKAL) is an optimization model. The focus is on the electricity and renewable components of these models which exchange price and quantity information with separate fuel supply and demand modules. All of the models use a bottom-up approach with technology options represented explicitly.

By pursuing a multi-faceted approach (optimization, detailed simulation and simple simulation), some of the shortcomings of the existing models can be addressed. First, the design described below allows the user to easily compare the output of a regional electricity forecast using all three of the solution techniques. Each technique can be prioritized by EIA and the system evolved as EIA gains more experience with the driving data sets. Second, the design has an intrinsic focus on the development of "smart" sub-regions avoiding "optimizing" regions that have serious geographical and demographic challenges (a significant number of islands, either literally or only in the abstract). Our design segments regional demand in a way that makes sense with regards to the technology slate, resource limits and levels of demand. Finally, the design places emphasis on making it more transparent and easy to test alternatives, displaying for the user the resultant "modeling" objects which collect together all the essential key data passed to the selected solution technique as well as the model results.

The design presented brings together the flexibility yet assists the user in managing the choices. This key feature will make the difference in the long-term sustainability of the model.

[1] For a more complete description see “World Electricity Models: Final Report,” prepared for EIA by OnLocation, Inc., August 8, 2014.

2 Input Requirements

As is true for the existing WEPS+ model, the new IEMM will require information that extends beyond the limited data passed from the other models. This section provides a brief discussion of the current data feeds of the electricity model and provides an outline of additional data that will help the new model (IEMM) better capture sub-regional conditions that will allow a richer forecast and more realistic estimate of future electricity supplies around the world. Section 3 will provide a more detailed assessment of the data needs and their potential sources. A key element of our design is the fundamental conviction that this additional data can improve the model's ability to forecast future generating capacity deployment and associated generation and fuel consumption.

2.1 Passed from Other Modules

Data coming from the other WEPS+ models (via a binary restart file, as in NEMS) include fuel prices by type, limited macroeconomic data, and sectoral electricity demands.

2.1.1 Regional Sector Demands

Demands are by year and by the 16 WEPS regions for residential, commercial, industrial and transportation

QELRS	Electricity Consumption - Residential
QELCM	Electricity Consumption - Commercial
QELIN	Electricity Consumption - Industrial
QELTR	Electricity Consumption - Transportation

2.1.2 Regional Fuel Prices

Fuel prices by year and each 16 WEPS region are delivered prices to the utility sector for distillate, residual, natural gas, coal and nuclear.

PDSPG	Distillate
PRSPG	Residual
PNGPG	Natural Gas
PCLPG	Coal
PNUPG	Nuclear

2.1.3 Macroeconomic Data

Macroeconomic data available to the electricity model by year and each 16 WEPS region include GDP, Population, Gross Output, GDP Deflator, Energy Price Deflator, Disposable Income, Working Age Population, and Employment.

GDP_PPP	Real PPP GDP over the forecast period in billions of year 2010 dollars
GDP_MER	Real MER GDP over the forecast period in billions of year 2010 dollars
POP	Population over the forecast period in millions of people
MacGO_PPP	Gross output over the forecast period for 20 industries in PPP billions of dollars
MacGDPD	GDP deflator, 2010 = 1.0
MacPrcD	Energy price deflator, 1982 = 1.0; not used in any model.
MacGO_MER	Gross output over the forecast period for 20 industries in MER billions of dollars
DisInc_PPP	Real PPP disposable income over the forecast period in billions of 2010 dollars
DisInc_MER	Real MER disposable income over the forecast period in billions of 2010 dollars
POP_Work	Working age population over the forecast period in millions of people
Employment	Employment over the forecast period in millions of people

2.2 Data Coming from Exogenous Input Files

The existing electricity model relies on considerable additional exogenous inputs to the international electricity modeling. These inputs characterize the existing and new power-generating technology as well as certain load and electricity pricing data. The inputs include[2]:

[2] Current Electricity Model Files:

- PG_UDI.xls (existing mw capacity by region, technology type, year)
- PG_TECHFILE.xls (new capacity costs and performance characteristics, exogenous learning parameters, capacity factors for new renewables, heat rate learning factors, optimism factors, capital costs learning factors).
- PG_INPUTS.xls (discount rates, logit coefficient, effectiveness of CCS factor, Electricity End Use Cost Factors, Reserve Margins, Fuel Price Factors, weights for fuel calibration, load shapes by region/sector/segment, T&D losses,)
- PG_NUCLEAR_PROJECTIONS.xls (forecast nuclear capacity and generation),
- PG_PLANNED_BUILDS.xls (planned builds and planned retirements),

- Existing power plant data (aggregate capacity by type, average O&M costs, heat rates and availability by technology type)
- New plant costs and performance characteristics (capital and O&M costs, heat rates, availability)
- Other assumptions (like transmission and distribution (T&D) losses, learning, etc.)
- Load data
- Electricity price information

Capacity data comes from UDI/Platts, *World Electric Power Plants Data Base*, that contains capacity data by unit and country, but does not include any cost or performance characteristics, so assumptions are made for these parameters. The model also relies on a number of parameters and inputs that are based on the analyst's judgment.

2.3 Evaluation of Existing Data Sources

In order to improve the model's ability to forecast future generating capacity and use, additional data is sought to allow a better characterization of the conditions 'on the ground' that would likely influence future decisions to add capacity and determine its use. The primary role of the additional data is to allow the modeler to match the conditions in the sub-regions with the technology slate most suited for evaluation in serving that area's electricity needs.

Section 3 describes in more detail the additional data that will assist in this richer characterization of the sub-regions demand density and economic and physical conditions (e.g. endowments of conventional fuels and renewable energy resource potential). Below are described in broad categories, the data and recommended data sources. For example, the following is sought:

- Additional macroeconomic data at the country level will be needed to help disaggregate WEPS+ regions
- Additional renewable and conventional fuel resource information
- Additional information to inform pricing (like subsidies and taxes)
- Actual data on existing and new plant costs and performance characteristics by country
- Information to help construct load shapes
- Data to help provide a distribution of fuel prices (rather than a single point)

- PG_RENEWABLES.xls (planned renewable capacity, renewable capacity targets, renewable capacity ceilings, forced generation)
- PG_REGIONAL_MULTIPLIERS.xls (Capital/Fixed O&M/Efficiency multipliers)

3 Data Assessment and Sources

In this section we step through the broad categories of information that can be used to enhance the characterization of the conditions in each sub-region and thereby improve the forecast of the future deployment of generating capacity to meet the sub-region's electrical needs and recognize that the data being gathered is at the country level. Our modeling approach will take the aggregate WEPS regional data and factor it into smaller sub-regions that can be addressed by one of our three solution algorithms (See Section 4 for a detailed discussion of each.) Our front-end solution will allow the user to select from a variety of economic and demographic factors to segregate the regional demand into smaller but similar shares of demand. When this is matched up to the technology slate appropriate to the sub-regional/sub-share, the solution approach best suited to address the capacity expansion, utilization and pricing is matched up.

This section aims to demonstrate that the additional data useful in aiding the detailed characterization of sub-regional electricity demands *is available* and can be assembled into a user-friendly front end that will allow the electricity analysis/modeler to improve on the electricity forecasting in WEPS.

3.1 Economic and Demographic Data

Economic and demographic data will be used, along with other data, to help characterize regions and to inform in the forecasting of regional- and sector-level electricity demand. Macroeconomic indices, such as gross domestic product (GDP), along with certain demographic data such as population density and the GINI coefficient, can be associated with electricity demand, especially in newly industrialized and developing economies. Specifically, the most useful macro-economic data will include:

- GDP by country
- GDP growth rate by country
- GDP per capita by country

Demographic data most useful will include:

- GINI coefficient (measure of income distribution) by country
- Population by region, country and city
- Population density by region, country and city
- Ratio of urban to rural population

- Percent of population with access to electricity

Illustrative macro economic data that can be acquired from international agencies such as the World Bank include that from:

International Monetary Fund (IMF), ***World Economic Outlook***, a comprehensive economic near-term survey of economic developments and policies in its member countries of developments in international financial markets, and of the global economic system. http://www.imf.org/external/ns/cs.aspx?id=29

World Bank, World Data Bank, ***World Development Indicators,*** a data bank of demographic data by country which includes the GINI coefficient for each country, the human development indicator and population metrics.
http://data.worldbank.org/data-catalog/world-development-indicators

City Mayors Statistics, ***The Largest Cities in the World by Population, Population Density and Area,*** city-level tables of demographic data including many metrics based on population.
http://www.citymayors.com/statistics/largest-cities-mayors-1.html

Central Intelligence Agency, ***The World Factbook,*** produced for US policymakers and coordinated throughout the US Intelligence Community, marshals facts on every country, dependency, and geographic entity in the world.
https://www.cia.gov/library/publications/the-world-factbook/

Trading Economics, *Forecasts/Economic Indicators,* is a publically available set of country-level forecasts for a wide variety of economic and demographic forecasts including GDP, credit ratings, tax rates and populations. http://www.tradingeconomics.com/forecasts

Shell, *New Lenses on Future Cities,* is a publically available report on projected growth of cities to 2050. http://www.shell.com/global/future-energy/scenarios/future-cities.html

3.2 Investment Environment (Corporate, State Owned, Other)

The investment environment in a country or region is largely dictated by the general health of the economy and the potential for return on investment to be realized. In developed countries/regions with strong economies, cost of money (interest rates), competition in a sector and government policy will influence both levels of investment and the type of player (e.g. private industry, government entity, regulated utility). When investment in pursuit of a particular policy is desired or a policy goal is set, governments can introduce an incentive or penalty schemes (discussed in section 3.4.2) to further influence investment behavior.

In less-developed countries/regions with weak economies, the role of government and multi-national development organizations becomes the defining factor for investment. In addition to direct investment from these organizations, sponsored incentive programs such as subsidies, loans and loan guarantees can help reduce the risk profile of a country and encourage investment. The United Nations Conference on Trade and Development provides data and analysis on investment in developing nations used to inform and influence macro-economic policy in those nations.

3.3 Factors Influencing Load Shapes and Demand

The ability to represent electricity demand (in the form of load duration) varies across regions and countries. For countries and regions with a wealth of information on electricity demand, load shapes and load duration curves can be constructed in great detail. In these cases, the actual load shape data will be used. However, in many parts of the world this data is not available. In order to address the regions with a paucity of data, our approach constructs representative load archetypes that, when combined, can be used to represent load shapes. Relying on general information on electricity consumption by sector, coupled with other economic and demographic data, load shapes and load duration curves can be developed. Our approach would develop ten load shape archetypes (four residential, four commercial and two industrial archetypes). Clearly other constructs are possible and can be investigated during the prototype model development phase.

3.3.1 Load Shape Archetypes

The load shape archetypes represent electricity consumption over an average weekday, and are categorized as:

Residential
Urban – Developed Economy
Urban – Developing Economy
Rural – Developed Economy
Rural – Developing Economy
Commercial
Urban – Developed Economy
Urban – Developing Economy
Rural – Developed Economy
Rural – Developing Economy
Industrial
Light Industrial
Heavy Industrial.

Primary drivers that determine load shape were selected to define the archetypes. For residential and commercial users, their existence in a rural versus urban environment and the general level of economic development were the key drivers of load shape. For example, it is assumed that developing countries use cooling, such as air conditioning, with less intensity than in developed countries (normalizing for weather differences), muting their peak load. For industrial customers, the type of industry (heavy versus light) held greater influence over load shape than the proximity to urban or rural areas, or the local economy. Heavy industry has a more consistent but higher level of load (high, flat load shape) than light industry whose load shape more closely resembles commercial load.

3.3.2 Differentiating Load Shape Archetypes and Load Duration Curves

The load shape archetypes provide a basis from which load shapes can be customized based on a country's specific economic, demographic characteristics and fuel availability characteristics. Indices such as GDP, GINI coefficient (an indicator of income distribution) and cooling degree days will be used to refine the load shapes for a country. GDP tends to be directly correlated to electricity demand and the GINI coefficient provides some indication of the concentration of wealth. In countries with a high GINI coefficient, the electricity usage is more likely concentrated in cities and amongst a relatively small population. Cooling degree days directly impact the amount of electricity used in a country. Especially for developed countries, the higher the number of cooling degree days, the higher the electricity demand and the greater the "peakiness" of the load shape.

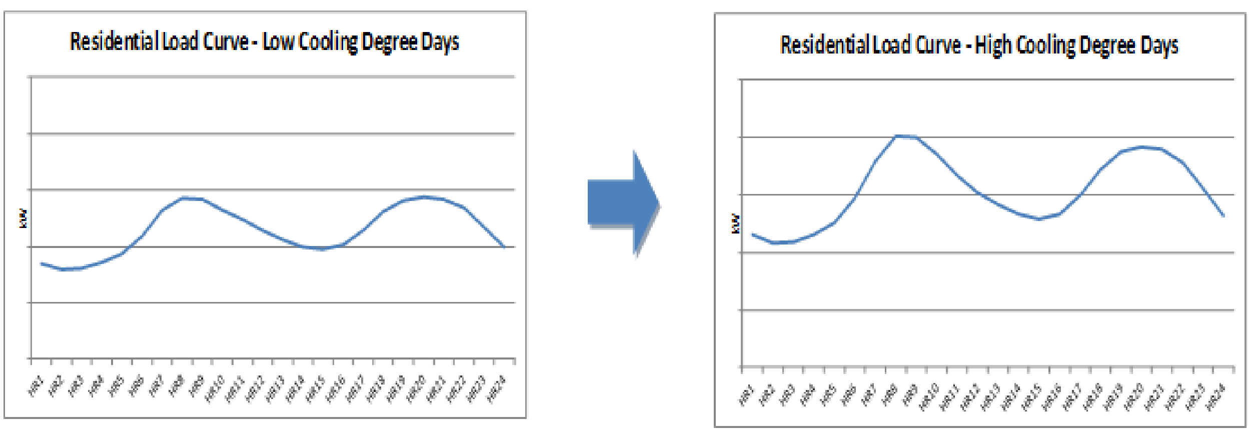


Figure 3-1: Alternative Residential Archetypes

For each country, data for electricity production by source can be used in combination with the load shapes developed from the set of archetypes to create a reasonable representation of an initial load duration curve. This information feeds the simulation model from which future capacity planning and dispatch are determined.

3.3.3 Data Sources

A broad set of economic and demographic data (including GDP, population, access to electricity, urban and rural population distribution, etc.) is available from the World Bank. Information regarding electricity load shape is available from local power pools and the Electric Power Research Institute (EPRI), which has made public a library of load shape curves based on geography, sector and building type.

World Bank, *World Data Bank*, *World Development Indicators,* a data bank of energy, economic and demographic data, by country which includes the GINI coefficient, GDP, population metrics, access to electricity and electricity source information among other items. http://databank.worldbank.org/data/home.aspx

Electric Power Research Institute, *Load Shape Library 2.0,* A load database and web-accessible repository of end-use and whole-premise data. The tool provides access to the best available end-use load data and whole-premise data by sector, region and building type in the USA. http://loadshape.epri.com/

3.4 Regional Policies

As one might expect, regional policies can have a significant impact on the development of electrical power in a region or country. These policies establish the economic and regulatory framework for investments in the power industry.

3.4.1 Corporate Income and Other Investment Related Taxes

The cost of capital invested by a firm is commonly calculated as the weighted average cost of capital (WACC).

$$WACC = E/V * Re + D/V * Rd * (1 – Tc)$$

where:

E = market value of the firm's equity

D = market value of the firm's debt

V = total value of the firm's financing (equity + debt)

Re = cost of equity

Rd = cost of debt

Tc = corporate tax rate

The corporate and other related taxes are therefore key in determining the cost of financing investment. Corporate tax rates (and tax rules that affect these rates) vary by country. Generally, lower corporate tax rates tend to encourage more investment. A country can selectively incentivize certain investment by reducing corporate tax rates and/or providing subsidies, or through other vehicles discussed in section 3.4.2. Corporate income tax data is publically available and is periodically printed from a number of sources including public accounting firms (KPMG; Ernst and Young; Price, Waterhouse, Cooper, etc.) and data aggregators such as Trading Economics.

3.4.2 Investment Incentives and Subsidies (positive or negative)

Investment can be stimulated through a number of incentive-based mechanisms. Some may be positive incentives such as tax holidays or subsidies for investment in certain technologies or regions, while others may be deterrents such as tax penalties on emissions. Positive incentives can be structured under a number of mechanisms including tax policy, subsidized loans or loan guarantees or rebates, to name a few. Incentives administered through tax policy can be broadly categorized as follows:

- Profit/income-based incentives: Reduced corporate income tax rate, tax holidays, loss carry forwards
- Capital investment-based: Accelerated depreciation, investment and reinvestment allowance, investment tax credits
- Sales-based: Income tax reductions based on total sales
- Import-based: Exemption from import duties on capital goods, equipment or raw materials
- Value-added-based: Income tax reductions or credits based on the net local content of outputs
- Labor-based: Reduction in social security contributions, deductions from taxable earnings based on number of employees3

Incentives for electrical system investment are frequently applied through one of the first four mechanisms in the list, however, incentives can take any form. For modeling purposes, these will be simplified into those that impact capital investments regardless of technology, and those that are targeted to specific technologies.

Negative incentives or policies can include penalties for not adhering to policies or missing policy targets, and options to purchase rights to exceed set standards (particularly for emissions), to name a few.

[3] United Nations Conference on Trade and Development (2000), *Tax Incentives and Foreign Direct Investment: A Global Survey, ASIT Advisory Series, No. 16;* United Nations, Geneva. UNCTAD/ITE/IPC/Misc 3.

Subsidies may also exist for both specific fossil and renewable resources and technologies, which affect the cost of new capacity and/or its production.

World Energy Council, *World Energy Resources: 2013 Survey*, is a publically-available document with information on country-specific renewable policies such as targets and subsidies.

REN21, *Renewables 2015 Global Status Report*, is a publically available report that contains information about policies to incentivize renewables such as feed-in tariffs.

IEA, *World Energy Outlook*, is available for purchase and contains detailed information about both fossil and renewable energy subsidies, in order to assess the impact of removing fossil fuel subsidies. IEA also maintains a ***Fossil Fuel Subsidies Database*** that is publically available.

Global Subsidies Initiatives (GSI) has some reports related to fossil fuel subsidy reform.

3.4.3 Taxes (VAT and other)

Taxes on goods and services (or consumption taxes - value added tax, "VAT", sales taxes and other, similar taxes) ultimately impact the price of electricity paid by the end user. Depending upon the type of tax, consumption taxes can also impact the effective price of imported and in some cases, exported goods (i.e. energy imports and exports). Data on consumption taxes is publically available and is tracked and published by a number of organizations including:

World Trade Organization, *World Tariff Profiles*, is a publically available data set of duties applied to goods in services by country and tariff type. https://www.wto.org/english/tratop_e/tariffs_e/tariff_data_e.htm

The World Bank, Taxes on Goods and Services, is a publically available data set that tracks current and historical tariff data. http://data.worldbank.org/indicator/GC.TAX.GSRV.VA.ZS

3.4.4 Discount Rates and Risk Premiums

Discount rates and risk premiums dictate the cost of investment. The discount rates are typically set by the central bank of a country and represent the cost of lending between banks. Risk premiums represent the relative likelihood of an investor receiving a certain return on investment and are applied to discount rates to determine the cost of money borrowed. For the IEMM, discount rates and risk premiums for a region or country will be largely influence by the sovereign credit rating of the country or member countries. Sovereign credit ratings are indicators of the relative economic and political risk of a country. Countries with good credit ratings (investment grade) have access to more funding sources at lower costs. For developing nations, a stronger credit rating also increases the attractiveness of foreign direct investment. Conversely, non investment grade countries (i.e. those with poor credit ratings) are limited in their access to funding sources and incur greater costs on debt. Three primary credit rating agencies (Standard and Poor, Moody, and Fitch) evaluate the credit worthiness of countries and periodically publish their findings.

Central bank discount rates are published by a number of sources including:

Trading Economics, *Country Interest Rates*, provides publically available data on several economic indices including country (central bank) interest rates as well as sovereign credit ratings from the three major credit agencies. http://www.tradingeconomics.com/forecast/interest-rate

Central Bank Rates, *Worldwide Central Bank Rates,* publically available listing of central bank rates. http://www.cbrates.com/

Central Intelligence Agency, *World Fact Book: Central Bank Discount Rate,* publically available publication of the annualized rate charged by a country's central bank to commercial, depository banks for loans to meet temporary shortages of funds. https://www.cia.gov/library/publications/the-world-factbook/rankorder/2207rank.html

3.5 Resources/Access to Fuels

The development of electricity in a country is directly affected by its access to resources and fuels used in generation. While some resources (such as fossil fuels) can be transported, others (such as solar and wind resources) cannot. Additionally, the cost of energy transportation (either fuels or electricity transmission) can be excessive and therefore not a viable option for some countries. The World Bank and other groups maintain data on the historical and current resources used for energy production by country. Several other organizations such as the Open Energy Information (http://en.openei.org/wiki/Main_Page) aggregate data from primary data sources on the global availability of renewable resources by country. Access to fuels and generating resources can expand or limit a country's electricity generation mix and potential accordingly.

3.5.1 Renewable Energy

The technical potential for new renewable capacity is dependent on access to renewable resources. Renewable resources for hydro, geothermal, biomass, wind and solar vary by country and region.

World Energy Council, *World Energy Resources: 2013 Survey*, is a publically available document with information on fossil and renewable resources for each member country. Renewable data include hydropower capability (theoretical, technical and economic) in GWh/year, peat land area, production and consumption, installed solar, wind and geothermal capacity and generation. It also includes information about country-specific renewable policies such as targets and other incentives.

REN21, *Renewables 2015 Global Status Report*, is a publically available report that contains information about the current status of renewable use, market and investment trends, and policies to incentivize renewables such as feed-in tariffs.

National Renewable Energy Lab, *International Resource Assessments and Maps*, have been developed for select countries for biomass, wind and solar resources. http://www.nrel.gov/international/global_energy.html.

Solar and Wind Energy Resource Assessment (SWERA) database is publically available. The resource maps can be used to visually identify areas rich in solar and wind resources within a country or region. The data products can help users estimate the size of a system that would be needed to meet specified loads and whether projects that incorporate these components are economically feasible.

International Renewable Energy Agency (IRENA) and **the Asian Development Bank** have several region-specific studies about renewable energy, TWh potential and costs, that are publically available.

Proceedings of the National Academy of Sciences (PNAS), *Global Potential for Wind-Generated Electricity*, includes TWh of potential wind energy for select countries and is publically available. **IRENA and DTU, *DTU Global Wind Atlas*,** has wind speed and power density maps for all countries.

3.5.2 Conventional Fuels (aka, coal, oil, natural gas, hydroelectric)

Some countries are endowed with rich fossil fuel reserves while others must rely on imports for fossil fuel generation. Some fuels may be available to a region as a whole (and provided at a price by the fuel supply modules) but would be prohibitively expensive to import into some of the countries within that region. Since the electricity model will be more disaggregated than the fuel supply modules, it is important to consider access at the country level where available.

World Energy Council, *World Energy Resources: 2013 Survey*, is a publically available document with information on fossil and renewable resources for each member country. Conventional fuels data include coal, oil, natural gas and uranium reserves and production, and some information on LNG trade and nuclear fuel cycle capability.

3.6 Technology Characterization and Regional Applicability

In addition to the characterization of the existing generating capacity around the world, there is a need for the characterization of new technology, and in particular new technology appropriate for large parts of the developing world. Further, in setting up the sub-regions in the model, knowing the extent of existing electrical transmission grids is constructive. This section identifies several sources of data to describe power generating technology in existence and available for new construction. Note, technology learning is discussed in section 4.3.1.

3.6.1 Initial Capacity and Characteristics

Some key requirements of an electricity model are to satisfy electricity demand, to determine fuel consumption needed for electricity generation, and to project future changes to the capacity (both new and retirements) and generation mix. It is important to have a detailed baseline of existing capacity and generation to start from. The baseline should include the current capacity mix (MW by technology and fuel type), operating characteristics such as heat rates for fossil plants and average capacity factors, and the variable and fixed operating costs of electricity production. Electricity imports and exports by country and transmission and distribution (T&D) losses are also needed to fill in gaps between electricity supply and demand.

Country-level information is generally available, whereas plant-level information is sparse. The following is a description of potential sources for this information:

UDI/Platts, *World Electric Power Plants Data Base,* is used by EIA for the current WEM model and contains detailed information about existing and planned power plants in every country. Data includes rated capacity (MW), plant status (operating, under construction, etc.), service dates, and technology type.

International Energy Agency (IEA), *World Energy Statistics and Balances 2015,* is a source that is used by EIA for the current WEM model. This set of publications has detailed information on most OECD member countries and more than 100 non-OECD countries. Data includes generation by fuel, fuel consumption, and electricity imports.

IEA, *Electricity Information 2015*, is available for purchase. Electricity *Information* has detailed information on 34 OECD member countries including electricity capacity and generation by fuel, fuel consumption in power plants, electricity imports by origin and exports by destination, T&D losses, fuel prices for electricity generation, and end-user prices of electricity. Average capacity factors and heat rates by fuel can be derived from this data.

World Bank, *World Development Indicators*, also has total electricity production and T&D losses by country. The information is based on the IEA data but is publically available.

United Nations, *Energy Statistics Yearbook*, includes electricity production, installed capacity and utilization by type for 235 countries. It is available for purchase.

CIA, *World Factbook*, provides electricity production, installed capacity by type, imports and exports for 267 world entities. Data is available online at https://www.cia.gov/library/publications/resources/the-world-factbook/.

3.6.2 Cost and Performance Data Characterizing New Generation Technology

EIA makes use of detailed cost and performance data characterizing generating technology in NEMS and to perform other analyses. However this data is focused on the USA. In order for an electricity model to project changes in the global capacity and generation mix, it is necessary to have estimated costs of technology choices, including how these costs vary by country or region. Variances in regional technology costs are a result of variations in labor costs and construction materials such as cement and steel, finance costs, O&M costs, fuel prices and other factors such as development experience and competitive markets. Sources for regional technology costs are described here.

World Energy Council, *World Energy Perspective: Cost of Energy Technologies*, is a joint study with Bloomberg New Energy Finance (BNEF) , that is publically available. Data include estimates by technology for capital and operating costs, capacity factor and levelized cost of electricity. Select countries and regions include US & Canada, Western Europe, China, India, Japan and Australia. **IEA, *World Energy Investment Outlook***, provides the assumptions by technology and country for investment costs, operation and maintenance costs and efficiencies used in their scenarios. The 2014 spreadsheet is publically available.
IEA, *Projected Costs of Generating Electricity,* is a joint report with the Nuclear Energy Agency and contains overnight capital, O&M and fuel costs as well as efficiency and LCOE by technology for select countries. The 2015 version is available for purchase but we have the 2010 version.
IHS CERA, *The IHS North American Power Capital Costs Index (PCCI)* and *The IHS European Power Capital Costs Index (EPCCI)*, tracks and forecasts the costs of building coal, gas, wind and

nuclear power plants associated with the construction of a portfolio of power generation plants in North America and Europe. Costs are indexed to the year 2000. This data is available for purchase.

4 Methodology Description

4.1 Model Objective(s)

Our modeling approach will take the aggregate WEPS regional data and factor it into smaller sub-regions that can be addressed by one of our three solution algorithms. Section 4.2 will describe our approach to breaking the regional demand into smaller yet abstract sub-regions. Our front-end solution will allow the user to select from a variety of economic and demographic factors to segregate the regional demand into smaller but similar shares of demand.

When this is matched up to the technology slate appropriate to the sub-regional/sub-share, the solution approach best-suited to address the capacity expansion, utilization and pricing is selected. In each instance, the objective of the solution approach is to determine the mix of new capacity that economically can serve the load, to determine the consumption of fuels (as categorized in WEPS) and to determine the price of electricity at the sector level for use in the demand models.

The reason for having three different solution approaches is to match the ’conditions on the ground‘ with the solution technique (if for instance an LP optimization is used to find the solution to an aggregate WEPS region, certain technologies only suited for smaller scale application will not appear economic.) Our three solution approaches are:

- Optimization with a linear program when significant data is available to characterize the existing and new capacity costs and performance parameters. It assumes a grid is in existence and power can flow relatively unimpeded across the region.
- Detailed simulation wherein sufficient data is available to describe load, but where there is limited information regarding the fuel costs, technology costs, etc.
- Simple simulation is intended to address a small portion of the load in remote or otherwise non-grid-connected conditions where a more restricted set of technologies is appropriate.

4.2 Regional Definition and Modeling Methodology Selection

Sub-regions/objects in our use case is intended to capture an abstract separation of the WEPS regional demands into logical ’buckets‘ of load that can be associated with connected grids, technology, income levels (and thereby influence estimates of load shapes), resources available and other factors that should be solved together to determine capacity, fuel consumption and prices. We use the term ’sub-region‘ to represent this abstract notion.

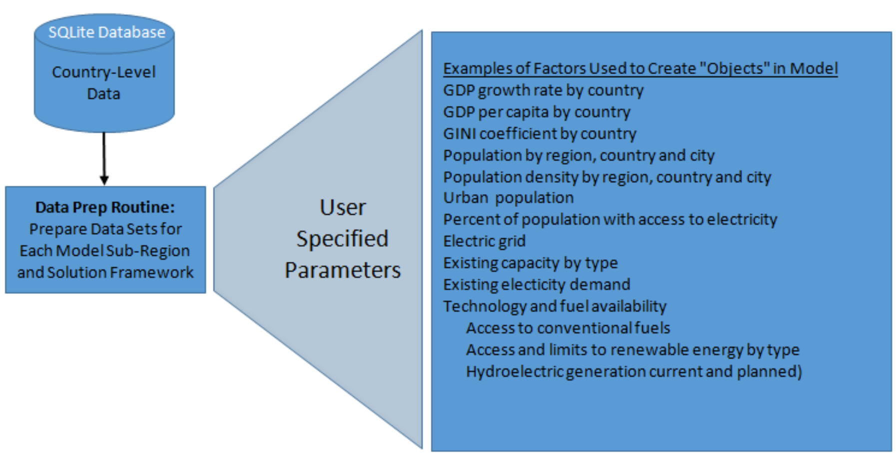
To illustrate the process of generating sub-regions/objects for solving, the following Table 4.1 shows a possible outcome for Africa (note this is only for illustrative purposes). In this case only a few factors were used to separate the aggregate Africa demand including the total size of the existing country-level generating system, the urban percentage (fraction of population living in urban areas), average income (GDP per capita) and the GINI coefficient (to skew income and electricity demand to the urban areas). In actual practice, the analyst will have a wide slate of data to aid in establishing breakpoints that produce reasonable sub-regions. Table 4.1 provides the result of this factoring of demand. If no other consideration was introduced, each of these shares of demand could be solved with the assigned algorithms and the results regarding consumption by fuel, electricity prices and technology details of capacity, generation and consumption aggregated back to the WEPS regional total.

Table 1: Summary of Illustrative Initial Shares of Africa Demand

Solution Technique Category	Starting Capacity (MW)	Number of Countries with Demand in Category	Regional Capacity Share
Optimize	89,257,004	9	64.3%
Detailed Simulation	41,460,996	18	29.8%
Simple Simulation	8,196,600	29	5.9%

However, as shown in Figure 4.1, additional considerations regarding technology and fuel availability (including access to and limits on renewable energy) would be used to further separate each of

Figure 4-1: User Interface to Establish Sub-regions



the shares. The introduction of these additional technology-/resource-limiting considerations lets us further separate the demand into packages that have consistent assumptions regarding grid connections, technology choices, renewable resources and conventional technologies at an appropriate scale. In the extreme, we could treat each individual country as its own isolated power market. In practice, demand from countries with similar assumptions regarding grid connections, technology choices and resource availability can be treated by the model as a single entity for which to solve. This is likely the case with the load identified as suitable for the Simple simulation and potentially the Detailed simulation. The Optimize category would likely be aggregated around grid connections.

Figure 4-2: Details Behind Illustration of Resultant Breakup of Africa

In Figure 4.2 above, each country in Africa has been assigned to one or more solution techniques. The green countries have been assigned to the Simple simulation, the orange countries to the Detailed simulation and the red countries assigned to two or more solution techniques. Note, the purple countries did not have sufficient information in our quickly assembled data set to allow for an assignment. This is an example of the user feedback that will be provided by the User Interface.

4.3 Common Elements

4.3.1 Technology Representation

4.3.1.1 Technology Suites

There is a very broad range of technologies used to generate electricity, ranging from small micro-turbines to large nuclear power stations. The natural scale of technologies and the types of fuels they consume determine their applicability to different situations. As described in Section 4.2, the regional definitions for the IEMM will be developed in part based on areas with commonalities in these dimensions. For example, areas with well-developed power grids that can accommodate large-scale technologies will be separated from areas that are more suitable for micro-grids due to dispersed demand and lack of transmissions access.

SMALLER SCALE GENERATING TECHNOLOGIES

- Reciprocating engines <20 Mw
- Gas Turbines <100 Mw
- Micro-turbines <250 Kw
- Fuel Cells <5 Mw
- Solar 1Kw+
- Wind turbines 200 Kw+

The slate of technology options will include the standard ones similar to those in most electricity models, including alternatives with carbon capture and storage, as well as a full range of renewable technologies. In addition, the insert above highlights some of the key smaller-scale generating technologies of potential importance to modeling a subset of the disaggregated regions that might be more rural and in the early stages of economic development. These technologies have not been typically included in the regional modeling due to their focus on small scale needs. While the amount of electricity these sub-regions will represent is not large relative to the total, it is important to recognize their potential role in the modeling framework. A recent General Electric report projects that distributed power capacity additions will increase to 200 GWs per year by 2020, up from 142 GW in 2012.

4.3.1.2 Technology Learning

The cost and performance of generation technologies has evolved considerably and is expected to continue to do so in the coming decades. Improvements stem from government and private research and development (R&D), spinoff effects from improvements in other technologies, and learning-by-doing that includes both incremental design improvements and manufacturing scale effects. R&D and spinoff from non-power technologies inherently cannot be modeled endogenously in an electricity model and so are best represented by user-specified trends over time. Learning-by-doing, on the other hand, is based on capacity deployments and so is best represented endogenously. Power technologies are global and many technologies share components, so the learning function will take into account both learning across regions as well as across technologies.

The capital cost of each technology will be characterized by its major component parts, such as turbine, heat recovery steam generator, gasifier, wind turbines, and the like. As a first pass, the component shares will be taken from the National Energy Modeling System's Electricity Market Module (NEMS/EMM). Total investment in each technology across all regions will be assessed for each simulation year. If desired, construction in some regions can be counted at less than its full value if that region is isolated from the global economy, and similarly the costs of technologies in those regions could be reduced by less than the learning rates applied to other regions. Typical learning curves will be applied with cost reductions occurring with each doubling of capacity. New components will have faster learning rates than those that are mature. Once the learning rates are calculated for each of the components, a weighted average learning factor is calculated for each technology.

4.3.2 Renewable Resources

For renewable energy technologies, a resource characterization is needed that reflects their availability throughout the year and allows their output profiles to be reflected in the technology competition. Resource supply curves with associated diurnal and seasonal profiles are a required input at the regional/sub-regional level. Fortunately, with the increasing interest in electricity modeling on an international regional basis, these data are becoming more available. The World Energy Council and International Renewable Energy Agency (IRENA) are among two potential sources (see Section 3.5 for more information). In addition, EIA has engaged the NREL to develop data on wind and solar PV availability. While information may not be as detailed as that available for the U.S., some rough estimates of the amount of capacity available at various levels of cost should be feasible to construct. The renewable resource data will be compiled and stored on a country-by-country basis allowing the IEMM to flexibly aggregate the information into alternative regional/sub-regional combinations.

4.4 Optimization Model Structure

4.4.1 Load Representation

The demand models of WEPS+ project annual electricity demands by economic sector for the 16 aggregate regions. As described earlier, these electricity demands will be characterized by archetype load shapes and aggregated by sub-regional demand. In regions where the optimization method is applied, the loads will be further processed into seasons and time slices.

4.4.2 Basic Optimization Model: Capacity Planning and Dispatch

4.4.2.1 Expectations - Investment Decision Framework

The optimization methodology will use a linear programming (LP) formulation to simultaneously determine short-term utilization of existing capacity to meet current demands and long-term capacity expansion decisions to meet future electricity demands. The dimensions of the investment decisions that are important include:

- Forward looking with respect to expectations of demands and fuel prices
- Ability to handle different risk profiles for regions and types of capacity (alternative technology or country discount rates can be used to capture relative risk)
- Ability to incorporate different capital asset lifetimes
- Ability to address comparative economics of long-lived assets
- Ability to address differing market or policy conditions

Because WEPS+ solves each of the sector models individually through the projection period and then iterates among them, electricity demand and fuel prices from prior solutions of the demand and fuel supply models can be used to inform the IEMM. Although foresight will be used, the LP will solve each forecast year, as in the EMM, rather than optimize across all future periods at one time. This allows the model to more reasonably reflect investor behavior in the light of changing economic and regulation conditions (i.e., see a carbon tax or RPS coming policy). This has the advantage of allowing for technology learning without creating a bias that new technology investments are made in order to make longer-term investments economic. The annual approach would also allow the models of WEPS+ to communicate on an annual basis rather than once per cycle if that proved to be a better solution in the future for convergence.

The proposed IEMM LP will use multiple time periods, much like the Liquid Fuels Market Model (LFMM) and Electricity Capacity Planning (ECP) module of the EMM in NEMS. Unlike the EMM, the dispatch of current capacity from the first period will be used and there will be no need for a separate dispatch routine. As in the ECP, we propose that the IEMM LP have three planning periods: the current year to capture dispatch decisions, the next year in which capacity build and retire decisions will be made, and a third period representing the remaining years in the planning horizon that informs the model regarding evolving economics and policies over the long-term and which allows for the model to delay expansion or retirement decisions. However, the number of periods will be specified in a flexible manner. The LP framework is designed to return solutions that:

- Meet specified electricity requirements in all seasons and time slices in all time periods.
- Meet capacity reserve requirements.
- Satisfy all other specified constraints.
- Determine the lowest price that can satisfy all electricity requirements while meeting investors' minimum rate of return investments requirements.

Because this is a multi-period model, allowing for the time value of money, i.e., discounting of the impacts of changes in the cost of key inputs and constraints needs to be included, all constraint right-hand sides and bounds (electricity demands, RPS constraints, etc.) in a given planning period are calculated as a net present value (NPV) average value. This average is computed as the NPV of the value in each year associated with a given projection period, divided by the net present value of one unit of product over the same time period. The first period costs will simply be those from the current year. In addition, all prices/costs in the LP formulation are calculated as the NPV nominal unit price discounted to the beginning of the full projection horizon. This value is calculated as the NPV of the values associated with each year of that planning horizon and then discounted to the first year of the projection horizon.

4.4.2.2 Index Definitions

$a \equiv$ *Index of generating technologies*
$b \equiv$ *Index of fuels*
$h \equiv$ *Index subset of hydroelectric generating technologies*
$m \equiv$ *Index of operating modes for electricity generation; minimum load requirements determine which technologies can operate in which modes*
$n \equiv$ *Index of electricity demand load slices*
$r \equiv$ *Index of external electricity demand regions*
$s \equiv$ *Index of seasons*
$x \equiv$ *Index of planning periods*

4.4.2.3 Column Definitions

These are the decision variables.

$B_{a,x} \equiv$ *New generating capacity builds in GW by generating technology (a) and from planning period (x) to subsequent periods*
$C_{a,x} \equiv$ *Commit vectors for the operation of existing capacity in GW are from period 1 through x*
$CC_x \equiv$ *Accounting variable representing the total capital cost from capacity expansion in planning period (x) in MM US$*

$CM_{r,x} \equiv$ *Carbon allowances bought from region (r) in planning period (x) in million metric tonnes (MMT)*

$CX_{r,x} \equiv$ *Carbon allowances sold to region (r) in planning period (x) in million metric tonnes (MMT)*

$FC_x \equiv$ *Accounting variable representing the total input fuel costs in planning period (x) in MM US$*

$F_{a,b,s,,x} \equiv$ *Purchase of fuel (b) by technology (a) in season (s) in planning period (x) in million MMBTU*

$FL_{b,s,,x} \equiv$ *Purchase of fuel (b) by in season (s) in planning period (x) in million MMBTU*

$M_{r,s,n,x} \equiv$ *Electricity imported from region (r) in season (s) in load slice (n) in time period (x) in GWh*

$O_{a,s,m,x} \equiv$ *Generating capacity of technology type (a) operated in season (s) and operating mode (m) in planning period (x) in GW*

$PO_{a,s,x} \equiv$ *Planned outage of generating technology type (a) in season (s) in planning period (x) in GW*

$R_x \equiv$ *Accounting variable representing the total revenue from electricity exports and the sale of carbon allowances in planning period (x) in MM US$*

$VC_x \equiv$ *Accounting variable representing the total variable operating costs in planning period (x) in MM US$*

$X_{r,s,n,x} \equiv$ *Electricity exported to region (r) in season (s) in load slice (n) in time period (x) in GWh*

4.4.2.4 Objective Function

The objective function seeks to minimize the net present value of total costs over all planning periods (x) in nominal dollars. This yields the lowest price that will support (cover the costs) all the current and future capacity and provides the required minimum rate of return on all capital investments. The total costs include fuel, other variable costs, and fixed costs, as well as any net revenue from imports and exports and any applicable carbon fees (if not already rolled into fuel prices). In the first period, available capacity is fixed, and the cost minimization determines the lowest cost operation of that capacity. This is the lowest price necessary to cover the cost of the marginal producer.

$$\text{Minimize} \sum_x \{FC_x + VC_x + CC_x - R_x\}$$

4.4.2.5 Description of Constraints

Fuel costs. For each planning period, the total fuel cost equals the sum of fuel purchases across all technologies, where fuel consumption is determined in the fuel balance rows as described below. Carbon fees or taxes are included here if applicable. Fuel and carbon prices reflect current year prices as well as projected future values. Fuel consumption and CO2 emission rates are characterized by technology as well as fuel type, so that emissions from technologies carbon capture and storage (CCS) can be properly computed.

$$FC_x - \sum_a \sum_b \sum_s (FP_{b,x} + E_{a,b} * CTAX_x) * F_{a,b,s,x} = 0$$

Where:

$FP_{b,x} \equiv$ *NPV nominal purchase price of fuel (b) in planning period (x) in US\$/MMBTU*
$E_{a,b} \equiv$ *CO2 emission rate per MMBTU of technology (a) using fuel (b) in tonnes/MMBTU*
$CTAX_x \equiv$ *NPV nominal price of carbon tax in US\$/tonne in planning period (x)*

The emissions' rates are technology- as well as fuel-specific in order to account for technologies with carbon capture and storage (CCS).

Variable Costs. The total variable operating cost equals the sum of variable operating costs over all electricity generating technologies, seasons, load slices, and operating modes.

$$VC_x - \sum_a \sum_s \sum_n \sum_m VOM_{a,x} * hours_{s,n} * DF_{m,n} * O_{a,s,m,x} = 0$$

Where:

$DF_{a,m,n} \equiv$ *deration factor for technology (a) in operating mode (m) that applies to load slice (n)*
$hours_{s,n} \equiv$ *hours per season (s) and load slice (n)*
$VOM_{a,x} \equiv$ *NPV nominal variable operating cost in US\$/MWh for generating technology (a) in planning period (x)*

The deration factor is tied to the operating mode and includes the effects of the forced outage rate, planned maintenance and load following.

Capital and fixed costs. The total fixed cost equals the sum of all capital plus fixed operating costs from capacity expansion over all generation technologies. Existing capacity has only fixed operating costs, while investment and fixed operating costs are attributed to new capacity.

$$CC_x - \sum_a XC_{a,x} * B_{a,x} - \sum_a FOM_{a,x} * C_{a,x} = 0$$

where:

$FOM_{a,x} \equiv$ *NPV of nominal fixed operation and maintenance costs for generating technology (a) in planning period (x)*
$XC_{a,x} \equiv$ *NPV of nominal annualized capacity expansion capital costs and fixed costs over the remaining years of the planning horizon for generating technology (a) in planning period (x)*

Trade Revenue. In those instances where sub-grids are represented with trade allowed between them, there is the possibility of electricity trade amongst the grids resulting in trade revenue. Total net trade revenue for a given planning period equals the sum of the revenue from the net export of electricity.

$$R_x + \sum_s \sum_n \sum_r \left[IPR_{r,s,n,x} * M_{r,s,n,x} - XPR_{r,s,n,x} * X_{r,s,n,x}\right] = 0$$

where:

$IPR_{r,s,n,x} \equiv$ *NPV nominal unit import price in USUS\$/kWh of electricity from region (r) in season (s) and load slice (n) in planning period (x)*
$XPR_{r,s,n,x} \equiv$ *NPV nominal unit export price in US\$/kWh of electricity to region (r) in season (s) and load slice (n) in planning period (x)*

Electricity load constraints. The operating generating capacity must satisfy native electricity demand plus net export demand in each season, vertical load slice, and planning period. Because each load segment represents a fixed time slice, the capacity and energy requirements are directly proportional and can be measured in either capacity or energy units. Committing capacity to each slice in GW units is equivalent to projecting electricity generation in kWh.

Technologies that are dispatchable (i.e. not intermittent renewables) can operate in one or more capacity factor modes, as in the ECP of the EMM and other electricity models, based on their required minimum number of hours of operation. In base mode, one unit of capacity would contribute energy to all load segments. If the same capacity is operated in intermediate mode, it would not contribute to those load segments with lower capacity requirements. This designation will prevent capacity types with long-start up times, like coal, from operating in a small number of hours.

For intermittent renewable capacity, a capacity factor is constructed for each load slice based on the coincidence of the renewable availability (i.e. when then sun shines or the wind blows) to load. Because the variable costs are essentially zero, the renewable capacity will be deployed up to its capacity factor in each time slice. Excess generation, which causes renewable curtailments, is allowed to occur if renewable generation is at high levels and generation must occur from other sources in order to meet the load in other time slices.

$$\sum_{a}\sum_{m} \mathrm{DF}_{a,m,n} * O_{a,s,m,x} + \sum_{r} \frac{[M_{r,s,n,x} - X_{r,s,n,x}]}{hours_{s,n}} - D_{s,n,x} \geq 0$$

where:

$DF_{a,m,n} \equiv$ *the deration factor for technology (a) in operating mode (m) that applies to load slice (n)*

$D_{s,n,x} \equiv$ *electricity demand in GW for season (s), load segment (n), and planning period (x) that has been grossed up for T&D losses*

$hours_{s,n} \equiv$ *hours per season (s) and load slice (n)*

Capacity constraints. The sum of the operating generation capacity and the planned seasonal outages must be less than the sum of the new and existing derated capacity for each generating technology in each season and planning period. In other words, more capacity cannot be committed to meet load than already exists or has been built in the period. Capacity that does not cover its fixed O&M costs across the planning periods becomes eligible to retire, but is only retired if the fixed costs are not covered in the first period as well. For example, a carbon emissions policy may make future operations of coal capacity no longer economic to operate, but it may still be able to operate enough to earn revenue in the current year to cover its fixed costs. In that case, the coal capacity would not retire in that year. In a subsequent year, if the policy becomes more stringent, the coal capacity may no longer be able to cover its costs in the current year as well and would be retired.

$$[B_{a,x} + C_{a,x}] - \sum_{m} O_{a,s,m,x} - PO_{a,s,x} \geq 0$$

BTU balance. The fuel use from operating generating capacity must be less than or equal to the amount of fuel purchased for each fuel in each season and planning period. The amount of fuel purchased depends on the hours of operation and the heat rate for each that capacity type.

$$\sum_{a(b)}\sum_{m}\sum_{n} DF_{a,m,n} * O_{a,s,m,x} * hours_{s,n} * heatrate_{a} - F_{a,b,s,x} \leq 0$$

where:

$a(b) \equiv$ *subset of generating technologies (a) that use fuel (b)*

$DF_{a,m,n} \equiv$ *deration factor for technology (a) in operating mode (m) that applies to load slice (n)*

$heatrate_{a} \equiv$ *heat rate for generating technology (a) in MMBTU per kWh*

$hours_{s,n} \equiv$ *hours per season (s) and load slice (n)*

Fuel use balance. The fuel balance row sums fuels across the relevant technologies in each season and planning period.

$$\sum_{a} F_{a,b,s,x} - FL_{b,s,x} \leq 0$$

Fossil fuel and biomass resource constraints. Fuel constraints are included so that limits on fuel supply can be represented. This is likely most applicable to biomass fuel. In this constraint, the fossil and biomass fuel use from operating generating capacity must be less than or equal to the amount of fuel purchased for each fuel in each season, load slice, and planning period. These fuel demands are passed to the supply models of WEPS+

$$FL_{b,s,x} - S_{b,s,x} \leq 0$$

where:
$S_{b,s,x} \equiv$ *Supply of resource in MMBTU*

Hydroelectric generation resource constraint. Hydroelectric generation is a special case because the resource is both energy- and capacity-constrained. The amount of generation from hydroelectric sources must be less than the maximum allowed in each season and planning period. The relationship of seasonal available hydroelectric generation to capacity may vary significantly by region.

$$\sum_{n} \sum_{a(h)} \sum_{m} DF_{a,m,n} * O_{h,s,m,x} * hours_{s,n} - H2O_{s,x} \leq 0$$

where:
$DF_{a,m,n} \equiv$ *deration factor for technology (a) in operating mode (m) that applies to load slice (n)*
$a(h) \equiv$ *Subset of generating technologies (a) that are hydroelectric*
$hours_{s,n} \equiv$ *hours per season (s) and load slice (n)*
$H2O_{s,x} \equiv$ *maximum amount of generation from hydroelectric plants for season (s) and planning period (x) in GWh*

Reserve requirement. The reserve requirement ensures that there is a sufficient amount of capacity to reliably meet peak demand. A simple reserve margin percentage will be applied to peak demand to determine the level of the requirement. If there is insufficient existing capacity, then new capacity will be built. The reserve allowance factor indicates the fraction of capacity for each technology type that can be counted towards meeting peak demand. For most technologies this is equivalent to the forced outage rate. For intermittent renewable technologies, such as wind and PV, the credit depends on the ability to provide power during times of peak demand and is assumed to decline as the share of intermittent generation increases in the region. The decline is designed to accommodate the variable nature of the resource and the need to protect against conditions in which it may not be available.

$$\sum_{a} RAF_a * [B_{a,x} + C_{a,x}] - \max_{s,n}[D_{n,x}] * \left(1 + \frac{RR_x}{100}\right) \geq 0$$

where:

$RAF_a \equiv$ *Reserve Allowance Factor indicating what fraction of capacity from generating technology (a) counts towards the reserve requirement*

$RR_x \equiv$ *Reserve requirement percentage indicating how much extra generating capacity above the peak load that must be held in reserve in planning period (x)*

Planned maintenance constraints. Planned maintenance can be addressed by further derating of the capacity. However, this has the effect of having the maintenance spread evenly across the seasons which will put a small bias on the need for new capacity[4]. If warranted, the model can be used to select the most economic season to do the maintenance. This section describes that internal determination of economic maintenance scheduling. Power plants are unavailable to generate due to planned maintenance as well as forced (unscheduled) outages. A simple approach would be to use historical trends to spread maintenance by season and then derate the capacity for both maintenance and forced outages. If the pattern of loads or other supplies, such as hydroelectric or other renewable generation, varies significantly by season, then it is better to allow the model to schedule the maintenance outages. An accounting constraint connects the seasons by ensuring that the hours of planned outage activity over all seasons are greater than or equal to the yearly planned outage hours required for each generating technology.

$$\omega_a * 8760 * [B_{a,x} + C_{a,x}] - \sum_{s} \sum_{n} hours_{s,n} * PO_{a,s,x} \leq 0$$

where:

$\omega_a \equiv$ *the fraction of yearly hours required for planned maintenance*

$hours_{s,n} \equiv$ *hours per season (s) and load slice (n)*

Renewable Portfolio Standards (RPS). A number of countries and sub-national regions, such as states in the U.S., have enacted requirements for a minimum level of renewable generation that is usually expressed as a percent of generation or sales. Here the assumption is that the requirement would be as a fraction of generation, but this could be modified to allow for sales if necessary. The technologies that are included in the RPS vary by country and some may

[4] The bias is a result of derating capacity in the peak season where normally one would not do maintenance. Thus, the need to fill in this gap results in a bias towards additional capacity.

include other 'clean' energy sources, either now or in the future. Therefore each technology will be given an RPS score that indicates its eligibility for the RPS in that region. In the LP, the constraint is expressed as the fraction of generation capacity from renewable sources must be greater or equal to the RPS requirement in each planning period (x).

$$\sum_{a}\sum_{m}\sum_{s}\mu_a * DF_{a,m,n} * \; O_{a,s,m,x} - \frac{RPS_x}{100} * \sum_{s}\sum_{\mathrm{n}} \mathrm{hours_s} * \mathrm{D}_{s,n,x} \geq 0$$

where:

μ_a ≡ *RPS score for generating technology (a)*

$DF_{a,m,n}$ ≡ *deration factor for technology (a) in operating mode (m) that applies to load slice (n)*

RPS_x ≡ *the percentage RPS requirement on capacity in planning period (x)*

$D_{s,n,x}$ ≡ *the electricity demand in GW for season (s), load segment (n), and planning period (x) that has been grossed up for T&D losses*

Carbon emission constraints. At present, various countries are pursuing a wide range of policies to reduce emissions of greenhouse gases, and in particular CO_2 emissions. Taxing carbon emissions or capping emissions and creating tradable allowances are two policies directly targeting reductions that have been enacted or which are under consideration in several countries. In the IEMM the easiest approach to represent is carbon emission penalties through taxes that can either be applied to fuel prices directly or incorporated in the IEMM operating costs. Implementation of a single-region emission cap in the power sector is also straightforward, while the WEPS+ framework of sectoral models makes economy-wide caps more difficult to represent. With a power sector cap, total CO_2 emissions must be below the emissions cap requirement in each planning period (x). This constraint is not applicable in a carbon tax case.

$$\sum_{a}\sum_{b}\sum_{s} E_b * F_{a,b,s,x} - CAP_x \leq 0$$

where:

CAP_x ≡ *carbon cap in MMT in planning period (x)*

E_b ≡ *CO2 emission rate per MMBTU of fuel (b) in tonnes/MMBTU*

If the policy to be modeled is a cap-and-trade CO_2 reduction policy across multiple regions, there are several possible methodologies. The most direct way would be to combine all the regions under the cap-and-trade policy into a single LP. Then the carbon emission constraints would become:

Carbon emission constraints for all cap-and-trade regions (r) and planning periods (x)

$$\sum_{a}\sum_{b}\sum_{s} E_b * F_{a,r,b,s,x} - CAP_x - \sum_{\bar{r}\neq r} CM_{\bar{r},x} + \sum_{\bar{r}\neq r} CX_{\bar{r},x} \leq 0$$

where:

$CAP_x \equiv$ *Carbon cap in MMT in planning period (x)*
$E_b \equiv$ *CO2 emission rate per MMBTU of fuel (b) in tonnes/MMBTU*
$F_{a,b,r,s,x} \equiv$ *Purchase of fuel (b) by technology (a) in region (r) in season (s) in planning period (x) in million MMBTU*

The total CO_2 emissions in each region would need to be below the emissions cap requirement for each region (r) plus net emission allowances imported from other regions in each planning period (x). This constraint represents a carbon cap-and-trade applied only to the electricity sectors in the included regions.

4.4.3 Alternative Combined LP Solution Approaches

There is potential concern that combining several regional LPs in one single large LP could result in unacceptably large solution times. Further, many of the regions will be assigned to solution techniques other than the LP. Additionally, this combination would be contrary to the proposed IEMM modular solution approach of solving each region in parallel. Fortunately there are decomposition algorithms for solving large-scale linear programs with the special constraint structure of the above combined LP. One common decomposition approach is called the Dantzig-Wolfe decomposition. This special constraint structure is comprised of a set of independent submatrices (F_i) which are connected only by a set of 'coupling' constraints (D_i). In the combined LP, the independent submatrices are the individual regional LP matrices and the coupling constraints are the new carbon emission constraints described in the previous paragraph.

The general idea behind the Dantzig-Wolfe decomposition is that the original large problem is reformulated into a reduced master problem and *n* sub-problems which may be solved in parallel (see Figure 4-3). This rests on the principle that a convex polyhedron may be described as a convex combination of its extreme points and rays alone.[5] Each column in the new master program represents a solution to one of the sub-problems and the master program enforces that the coupling constraints are satisfied. The master program then requests additional solutions (columns) from the sub-problems, such that the overall objective to the original linear program is improved and the process is repeated.

---

[5] This result is called the Minkowski's Representation Theorem.

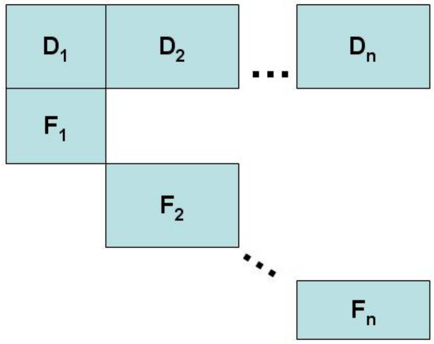


Figure 4-3: Decomposition Illustration

The general solution algorithm is:

Initialize the reduced master problem with a feasible set of columns

1. Solve the reduced master problem
2. Use the duals from the master problem to formulate and solve the sub-problems in parallel
3. If the columns suggested by the sub-problem solutions have a negative reduced cost, add them to the reduced master problem
4. Repeat steps 2-4 until there are no more columns generated with a negative reduced cost

This approach has the advantages that the sub-problems may be solved in parallel and that it allows the modular data structure to be maintained.

This methodology could be used to carbon cap-and-trade policies with multiple region participation in the same program or for inter-regional electricity trading.

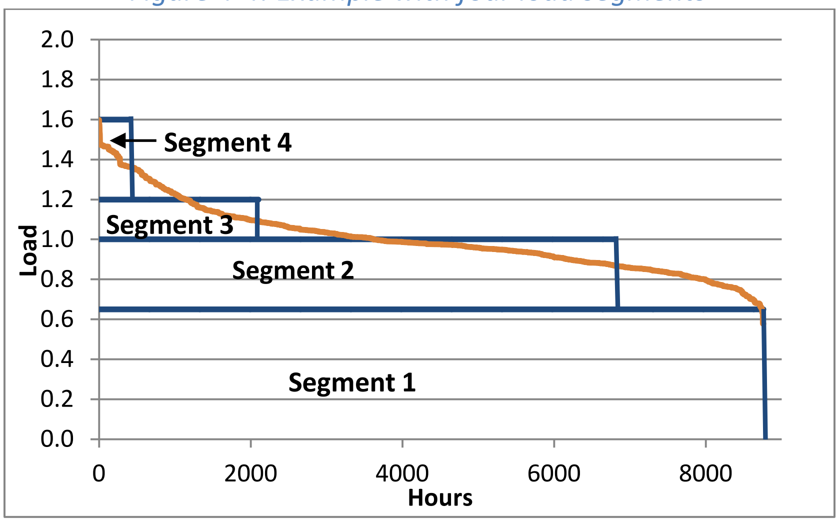
4.5 Detailed Simulation Model Structure

As described previously, we propose considering two levels of simulation modeling as well as the optimization method. The Detailed simulation approach described in this section is proposed for areas in which adequate data is available (and potentially a fairly developed grid is in place). These regions will represent sub-regions based on the resources available, relatively significant electric loads, appropriate technologies and other factors. Regions with very little data and minimal grid infrastructure will be modeled using a Simple simulation method described later.

4.5.1 Load Representation

Annual electricity demand will be divided into several load segments, representing periods of varying levels of demand and hours, which will be used in the capacity expansion and dispatch. This is similar to many other global simulation models, where two to six load segments have been used. The number of blocks will be flexibly defined in the IEMM, but we imagine that four to five will be sufficient to represent different generating modes while not over-interpreting the estimated load shapes. As an example, the baseload segment would represent 8,760 hours, while the peak would represent perhaps 400 or so hours. In addition, an annual peak demand will be projected. In both simulation methods, as in the optimization method, demand from the other WEPS+ models are grossed up for losses.

Figure 4-4: Example with four load segments



4.5.2 Capacity Planning: New Capacity and Retirements

Capacity planning consists of selecting a mix capacity that meets projected electricity demand and peak load requirements. New and existing capacity will compete within each of the load segments. Expansion will occur when the current year existing capacity is less than peak demand plus a reserve margin, or when new capacity is more cost-effective than existing capacity. In a business-as-usual scenario it is unlikely that much, if any, existing capacity will retire, but in a carbon price scenario, existing coal capacity may be sufficiently penalized to make some portion of existing capacity no longer economic. In addition, specified minimum retirement rates by technology type can be used to retire a portion of existing capacity. For example, nuclear power plants generally have fixed operating licenses, although assumptions can be made about their extension.

Capacity that is counted towards meeting reserves would vary by technology type depending on typical forced outage rates by technology. Intermittent renewable capacity would be represented by a dynamic capacity credit as a function of the share of renewable generation in the region in the current year.

4.5.3 Technology Shares

A set of eligible technologies will be established for each region's new power supply options. Projected future supplies will be selected from these technologies based on economic and other criteria. For example, natural gas supplies may not available in some regions or countries may have stated moratoriums on new coal or nuclear power, while other regions may have a full suite of options. Large-scale hydroelectric power development will be treated exogenously because the drivers for development are broader than just the provision of electricity.

The shares of technologies used to meet new supplies will be based on a logistic function. The representation of each technology as having a distribution of costs rather than just a point value is intended to capture the range of fuel costs, variation in the installed cost based on location, and variations in the cost of financing. There are a few standard ways logistic equations are used for determining market shares. One method is to apply a coefficient to each of the technology attributes, such as capital cost and operating cost, to compute a 'utility' of each technology option. Then the shares are computed based on:

$$\mathrm{share}_i = \frac{\mathrm{e}^{\mathrm{U}_i}}{\sum_i \mathrm{e}^{\mathrm{U_i}}}$$

where:

U_i = Beta1 * capital cost_i + Beta2 * operating cost_i + ... + Beta_N * attribute_N_i

This is the method used for efficiency technologies in the residential model and light-duty vehicle types in the transportation model of NEMS. The ratio of the coefficients for capital and annual operating costs can be interpreted as a simple payback period. Other attributes that represent either a general technology bias or specific other features (for example, acceleration capabilities for vehicles or intermittency for generation technologies) can be easily included. Another methodology is to create a single lifecycle cost metric which can be multiplied by a coefficient and treated as the utility in the equation above. In the power context, this lifecycle cost could be the levelized cost of energy (LCOE) or the NPV of capital and operating costs.

An alternative formulation is to compute the shares using the equation:

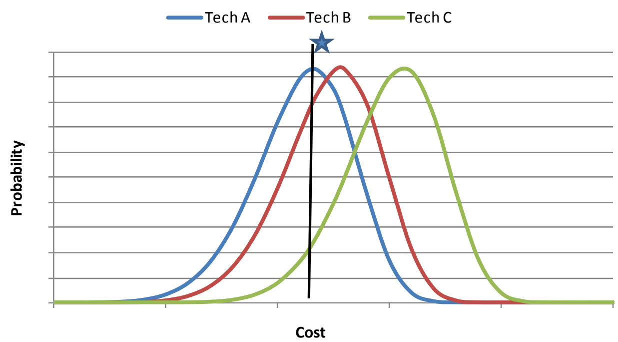
$$share_i = \frac{{E_i}^{-\lambda}}{\sum_i {E_i}^{-\lambda}}$$

where:

E_i represents the economic indicator of LCOE or NPV of expenditures for each technology *i*.

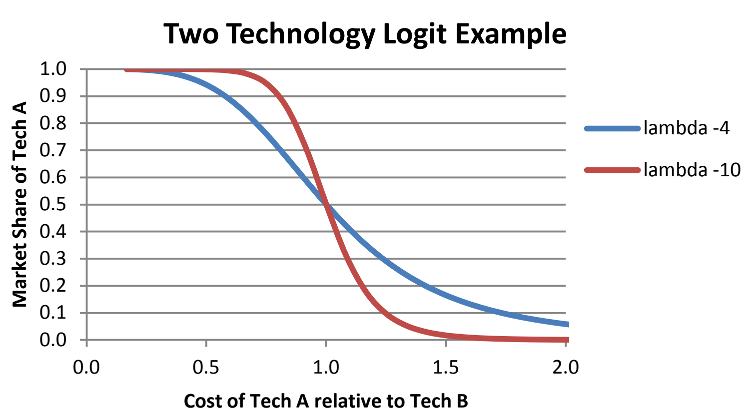
Figure 4-5 illustrates a logit function with three technologies. Although technology A has the lowest mean cost, market shares would also be allocated to technologies B and C.

Figure 4-5: Example of logit distribution



The lambda parameter indicates the width of the distribution. If the value is large, then the cost distributions are narrow, and the lowest cost option will gain most of the market share. If the value is small, then the shares will be more equal across the options. Many of the electricity simulation models we reviewed use LCOE as the metric in the logit, sometimes with other factors included as either multiplicative or added into the total cost in their logistic formulations.

Figure 4-6: Two technology logit example with lambda



For the IEMM Detailed simulation methodology, we propose to compute an aggregate cost for each technology and to use the lambda equation form. A multiplier is introduced to represent non-cost factors that can be set to calibrate the results to recent historical experience.

$$share_{s,i} = \frac{mult * E_{s,i}{}^{-\lambda}}{\sum_i mult * E_{s,i}{}^{-\lambda}}$$

where:

mult $_T$ represents the influence of non-cost factors,
s represents the load segment, and *i* is the technology

This competition is performed for each of the several load segments using number of hours that segment[6]. Only the technologies applicable to the segment would be included. Alternatively, the exclusion could be represented by including cost multipliers that reflect the impact of ramping on individual technology types. For example, new nuclear power would likely only compete in the base segment and only gas combustion turbines, diesel generators, and PV would compete in the segment representing high levels of demand that occur infrequently. A unique feature of our methodology for the Detailed simulation is that it includes existing as well as new capacity in the competition within each segment, rather than only perform a competition for a predetermined amount of new capacity. The cost metric is the NPV of total cost in nominal dollars for a unit of capacity that is comprised of three main components: fuel, non-fuel variable and fixed costs.

$$\mathrm{E_{i,s}} = FC_{i,s} + VC_{i,s} + CC_{i,s}$$

Where FC is the fuel cost, VC the variable non-fuel operation costs and CC is the fixed cost for technology *I* for segment *s*.

The fuel costs are based on the NPV of costs over time that take into account future fuel price escalation as well as incorporate any carbon emission penalties (if not already incorporated in fuel costs by the rest of WEPS+). Because WEPS+ runs the models in parallel and cycles, expectations of fuel prices can be taken from the prior cycle for the future years.

$$\mathrm{FC_{i,s}} = \mathrm{NPV} \sum_{y,i} (FP_{y,i} * heatrate_i + CO2P_{y,i} * \frac{\mathrm{CO2}}{\mathrm{BTU}}) * (\mathrm{hours_s})$$

The non-fuel variable costs consist of variable O&M cost and are expressed as the NPV in nominal dollars of the future cost stream for using technology *i* in that load segment.

$$\mathrm{VC_{i,s}} = \mathrm{NPV} \sum_{y,i} (VOM_{y,i}) * \mathrm{hours_s}$$

[6] Alternatively, the costs can be levelized costs per kWh. This will yield the same relative costs between technologies and hence market shares.

For existing capacity, the fixed cost component only includes the fixed O&M costs. For new capacity, fixed costs also include a capital cost that will reflect costs appropriate for the sizes of the technologies deployed. The fixed charge rates are also technology specific and take into account different tax treatment, as well as construction times that impact the financing costs.

For existing capacity:

$$\mathrm{CC}_{i,s} = \mathrm{NPV} \sum_{y,i} (FOM_{y,i}) * \mathrm{availability}_i$$

For new capacity:

$$\mathrm{CC}_{i,s} = \mathrm{NPV} \sum_{y,i} (capcost_i * FCR_i - capcredit_i + FOM_{y,i}) * \mathrm{availability}_i$$

New capacity is also given a capacity credit that is based on the fixed costs of a new combustion turbine (or diesel generator) and depends on whether new capacity is needed due to load growth, as well as the potential contribution of technology *i* to peak demand. This recognizes that new capacity has value for contributing to capacity requirements as well as providing energy in that segment. The peak contribution for intermittent renewable technologies will decline as its regional share increases.

If existing capacity < peak load * reserve margin,

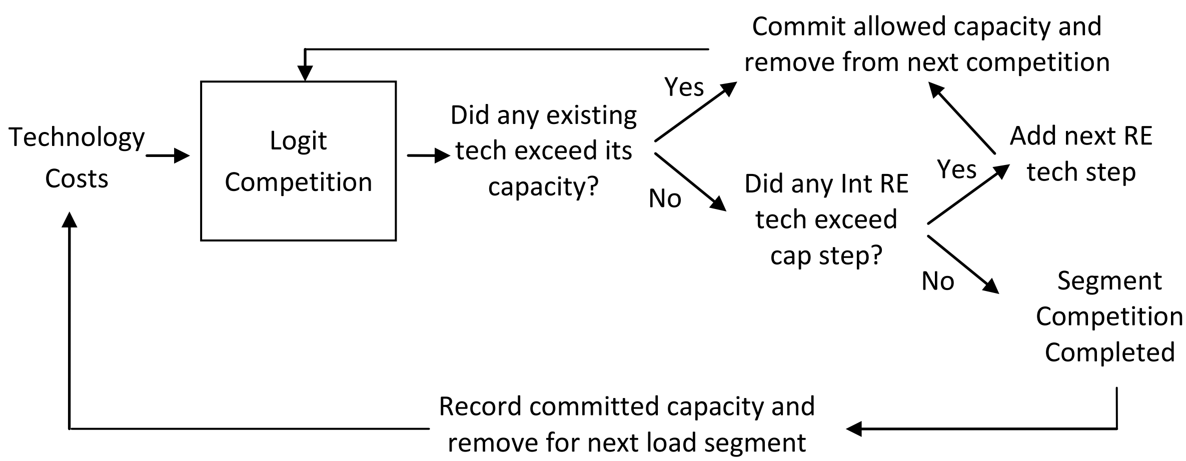
$$\mathrm{capcredit}_i = CC_{CT,s} * \mathrm{peakcontribution}_I$$

Otherwise, Capacity credit = 0

The competition among technologies is first performed for the baseload segment. In a reference case, the existing capacity will likely take most of the share, except what might be needed to meet new demand. The logit market shares are computed iteratively, with a check on the amount of existing capacity available. Because existing nuclear and hydroelectric generation will always be at the lowest cost, the amount of these capacities is subtracted from the load segment first. If the first market shares of the remaining technologies yield a share for an existing technology that is greater than the existing amount, the full amount of that capacity type is selected, and another competition is performed among the remaining technologies. For each iteration, any capacity that receives a small share less than a user-specified minimum size for that technology is ignored. The iteration continues until the full load of the segment has been satisfied. The amount of new capacity that can be selected is unlimited (see following section for special considerations for intermittent renewable capacity).

Once capacity has been selected for this first segment, existing capacity is decremented by the amount deployed in that segment. This lower capacity is then available for selection in the next load segment. The competition will lead to different market shares in this segment as the number of operating hours is lower. Again the logit is used but limited by iterating to selecting quantities of existing capacity that are available. After the load in the segment is met, existing capacity it again decremented for the next segment. In the peak segment, only a limited set of technologies compete. Coal and gas combined cycle plants are not suitable to meeting this demand because they cannot ramp up for just the limited number of hours represented in the peak segment.

Figure 4-7: Flow of capacity expansion within a load segment



A final segment for peak capacity is evaluated. The capacity of this is equal to the reserve margin times the peak load with no energy. Therefore the competition is performed on the capital component only for existing turbine or diesel generator capacity that has not been used to meet load and new turbine or generator capacity.

Special considerations for renewable capacity

Intermittent renewables (primarily wind and PV) require special treatment because their availability depends on regional resources and varies significantly over the course of a year. As a result, the capacity factor for each load segment will be related to the coincidence of the availability and load. In the base segment that represents all hours in a year, the availability equals the annual average. In the other segments, the availability could be the same if the renewable energy generation is equally available all year, or its pattern is unknown. Alternatively, it might be higher in the peak segments if the renewable energy is more available when demand is highest, such as may occur with PV that is available in daylight periods.

Because PV generation is likely to be correlated to times of high demand, it has more value in the base segment than the average value across all the hours. To reflect this additional value, a credit is given to PV equivalent to the difference in the base segment price and the generation time-weighted price. The prior year's average costs for each segment could be used to establish this value.

The intermittency of wind and PV also imposes a cost to the system as its share increases and other generators must ramp up and down more quickly. These grid costs can be approximate through an ancillary cost penalty that increases as the wind or PV share of generation increases in the region.

Finally, as described elsewhere, good resource sites may be limited for wind, solar concentrated power, and geothermal power. This can be represented by a supply curve representing potential capacity at successively higher costs can be used. In the market share competition, once the capacity at a cost step is reached, it is removed from the competition and replaced with the capacity at the next cost step. Similar to the treatment of existing capacity, the market shares are then recomputed.

The sum of the capacity serving each of the segments determines the total capacity available. Because the model uses future demand and energy prices with perfect foresight, no explicit delay in power plant construction is represented. If some existing capacity is unused in any of the segments, it is kept for a few more years. After that, if it continues to not be selected, it is retired.

4.5.4 Dispatch

For the Detailed simulation, the capacity expansion determines the new capacity added and retired, and then those assets are dispatched. The first step is to allocate the intermittent renewable generation to each load segment. The rest of the dispatch is much like the generation planning, except the only variable costs are included in the cost metric and current rather than levelized future fuel and CO_2 prices are used.

For dispatch:

$$\mathrm{E_{i,s}} = \mathrm{NPV}\sum_{\mathrm{y,i}}(\,FP_{\mathrm{y,i}} * heatrate_i + CO2P_{y,i} * \frac{\mathrm{CO2}}{\mathrm{BTU}} = (\,VOM_{\mathrm{y,i}}) * (\mathrm{hours_s})$$

First, the baseload segment is dispatched using the logit competition. Each technology type is limited to its already selected new and existing capacity, and the logit is recomputed with the remaining capacity available as necessary. The same type of iteration is performed as for the capacity expansion.

4.6 Simple Simulation Model Structure

A simpler version of the simulation method will be used in regions where very little power sector data is available and the applicable generation technologies are limited. This technique will be applied to regions that are primarily comprised of rural areas of less developed countries without much existing power grid infrastructure. In these areas primarily small scale technologies apply, such as micro turbines, diesel generators, small-scale wind and PV.

In this approach, electricity demand will be characterized by kilowatt-hours *without* a specific load shape. Growth in demand each year will dictate the need for new capacity. This implicitly assumes that there is currently no surplus in generating capacity in these undeveloped regions.

In the Simple version, the mix of generation is a result of the technology competition with the new generation by technology type being added to the existing.

One competition will be performed among all the appropriate technologies. The hours used in computing the NPV of fuel and operating costs is the typical number of hours that each technology is expected to operate. This treats the entire load requirement as a single competition.

For the Simple simulation:

$$\mathrm{FC}_{\mathrm{i,s}} = \mathrm{NPV}\sum_{y,i}\left(FP_{y,i} * heatrate_i + CO2P_{y,i} * \frac{\mathrm{CO2}}{\mathrm{BTU}}\right) * (\mathrm{hours})$$

$$\mathrm{VC}_{\mathrm{i,s}} = \mathrm{NPV}\sum_{y,i}(VOM_{y,i}) * \mathrm{hours}$$

$$\mathrm{CC}_{i,s} = \mathrm{NPV}\sum_{y,i}(capcost_i * FCR_i + FOM_{y,i}) * \mathrm{availability}_i$$

The resulting shares will be applied to the increase in generation needs each year. Capacity will be imputed using typical capacity factors for reporting purposes.

4.7 Estimating Sector Prices

The IEMM must produce retail electricity prices by sector to pass to the demand models of WEPS+. These prices will be built up from generation prices, T&D adders, and taxes with a gross up to account for losses. Prices will be projected by the IEMM sub-regions and then aggregated to the 16 WEPS+ regions. While the fundamentals are similar, the methodology will vary somewhat between regions using the simulation and optimization approaches because different levels of information will be available in each.

Generation Component of Prices (Wholesale Prices)

Two methods will be used for determining generation prices: average cost pricing and marginal cost pricing. The average cost pricing approach will be applied in regions where the simulation approach for capacity planning and generation is used. In regions where optimization is used, the electricity pricing methodology, either average cost or marginal pricing, will be a user-specified option, and can include a weighted average blend of the two approaches where desired.

Average Cost Based

Average cost electricity pricing is used to establish retail rates in much of the world or can be assumed to approximate it where prices are set by governments. The generation component of average cost prices is a function of the value of current assets, capital expenditures for new capacity additions, and operating and fuel costs of generation. This methodology will be used for the simulation regions, as well as be an option for regions modeled through optimization.

In the Detailed simulation and optimization methods, the average production cost for a region can be computed as the sum of the annualized capital costs plus operating costs. The generation price is then the generation cost grossed up for losses.

$$\text{AvgGenCost} = \sum \frac{\text{CapCosts} + \text{OpCosts}}{\text{Generation}}$$

and

$$\text{AvgGenPrice} = \frac{\text{AvgGenCost}}{(1 - \text{loss factor})}$$

The capital component is the sum of an existing cost basis, the annuities for capacity built in prior years, plus the annuities for the capacity built in the current year. The existing capital component declines according to the fixed lifetime rates, but remains if the plant is economically retired. In other words, if a carbon policy leads to a premature retirement of fossil fuel capacity, the cost of the capacity remains in electricity rates until the end of its natural lifetime.

$\text{ExistBaseCost}_i\,(y) = \sum \text{ExistBaseCost}_i\,(y\text{-}1) * \text{decay rate}_i$

The decay rate equals one over the average lifetime if the starting age of plants is uniform. Regions where electricity demand is growing rapidly will have a lower decay rate.

$$\text{CapCosts} = \sum_i \text{CapAnnuities} + FOM_i * \text{GenCap}_i$$

where CapAnnuities have been accumulated as new capacity is built.

An initial capital component can be estimated based on the existing capacity mix and rough estimates of that cost of that capacity. Alternatively, assuming we can acquire detailed delivered cost of electricity and separate out the T&D adders and subsidy/added taxes, then we can estimate the capital component adder of existing generation.

Operating costs are comprised of fuel and non-fuel cost components. Fuel prices are the current year prices. In fuel cost calculation, the current heat rate will be used for each technology type in regions using the simulation approach. In the optimization regions, discrete vintages of capital will be tracked and the heat rate may vary by capacity vintage.

$$\text{OpCosts} = \sum_i (\, FP_i * \text{heatrate}_i + \text{VOM}_i) * \text{Gen}_i$$

In the regions that are simulated using the Simple simulation approach where capacity is not explicitly represented, generation prices will be approximated using a weighted average of the levelized NPV cost for all generation types as calculated each year. The generation weights would reflect the technology mix of generation that includes existing as well as new sources for each year.

Avg generation price = Avg levelized cost / Total Prod

Marginal Cost Based

Marginal cost pricing, which is used in some countries including parts of the U.S., reflects those markets where producers bid into a competitive market and all generators are paid the clearing price that matches supply to demand. Many of these markets include capacity as well as energy markets.

The optimization methodology for capacity planning and dispatch will produce information about the marginal cost of generation from the duals of the LP. The short-run marginal cost of generation will be by season and time period, while the marginal cost of capacity will be annual. The generation price would be computed as the load weighted average marginal energy price plus a capacity component. The capacity component is equal to the capacity dual times total capacity required divided by generation, where the total capacity required is equal to peak demand times the reserve margin.

T&D and Sectoral Markups

Given that there is little detailed information about the components of international electricity prices, we propose to use a relative simple approach for computing retail prices from generation prices. Historical retail prices will be used to construct simple markup from the average generation price, as computed above, to sectoral retail prices. Ideally this would be performed at the country or sub-national level. The markups could then be rolled up to the IEMM regions in a dynamic way that is consistent with the selected regional definitions. Where insufficient data on retail prices exist, markups will be constructed by using analogous regions. In regions where there is sufficient data available, the markup will be disaggregated into two components: cost of T&D and taxes/subsidies.

4.8 Aggregating to WEPS+ Regions

Once retail prices are established for the IEMM regions, average prices will be constructed for the WEPS+ 16 regions. These will be weighted averages using the same assumptions about demands in the IEMM regions relative to the WEPS+ regions. Similarly, the fuel consumption will be aggregated to pass to the other WEPS+ models. Capacity, generation and other electricity data will be aggregated at the WEPS+ region level and elsewhere reporting at the sub-region level to allow verification and validation.

5 Outputs

5.1 Passed back to other modules

Information passed back to the other WEPS+ models via the restart file include fuel consumption by fuel type, and electricity prices by sector. In addition, capacity and generation projections by fuel are also provided in the restart file for reporting in the standard IEO tables. Items provided in the restart file are available to GrafNEM.

5.2 Additional Electricity Reports

The current electricity model produces an excel output file that includes new and existing capacity and generation by fuel and technology. It is not a formatted report, but is available for independent processing. We anticipate providing a database of electricity data and creating some standard electricity reports from the IEMM that are described later in this CDR report section.

6 Running the Model

6.1 Standard WEPS Flow: Iteration to establish supply-demand balance

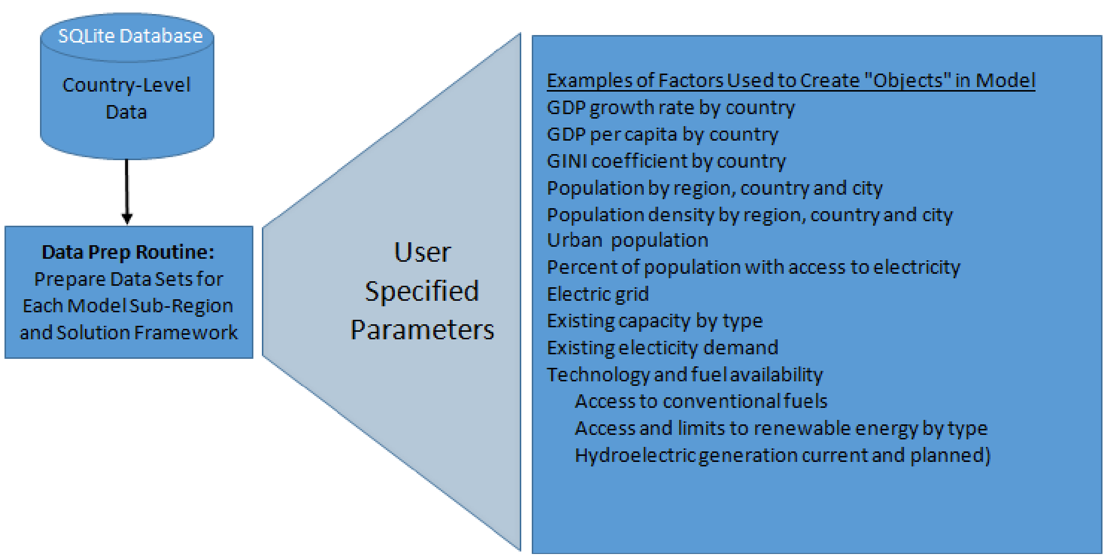
The IEO is generally run in integrated mode (with all models on), but models can also be run individually, pointing to a common restart file.

The process of making integrated runs will not change with the new IEMM. However, the process for electricity model developers to set up inputs and test the new IEMM will be completely different and is described below.

6.2 User Interface

While the raw form of many of the IEMM tables will remain as CVS, the plan is to implement a system that brings all of the data that is used by the IEMM (including the restart variables) into an SQLite database, and which will have a front end that allows the user to select parameters to be used to disaggregate the 16 IEO regions to whatever sub-regions as desired. The system will provide feedback to the user as to how many 'objects' are a result of their parameterization and the matching to resource and technology slates (other parameters could also be used to influence this choice). Once the user is satisfied that the sub-regional setup is acceptable, the program will produce the objects needed by the Python electricity model (see next section regarding model generation). The electricity model will then solve and provide a standard set of reports as well as an output SQLite database.

Figure 6-1: User Interface to Establish Sub-regions (repeated from above)



While not essential, we recommend the Front End include a GIS that provides feedback to the user as to the impacts of their choices when viewed at the WEPS regional level, or at the further disaggregated sub-regional level. Further, the user will be able to visually view the properties associated with each sub-regional object for validation and verification purposes. Finally, the GIS system can be used to facilitate the 'benchmarking' and associated imputation of missing or miss-scaled data. See Figure 6-2 for an example of additional feedback that can be presented to the user to assist in the definition of sub-regions. Again, this is only illustrative but it contrasts GINI (wealth concentration) with kWhPerCap (electric use).

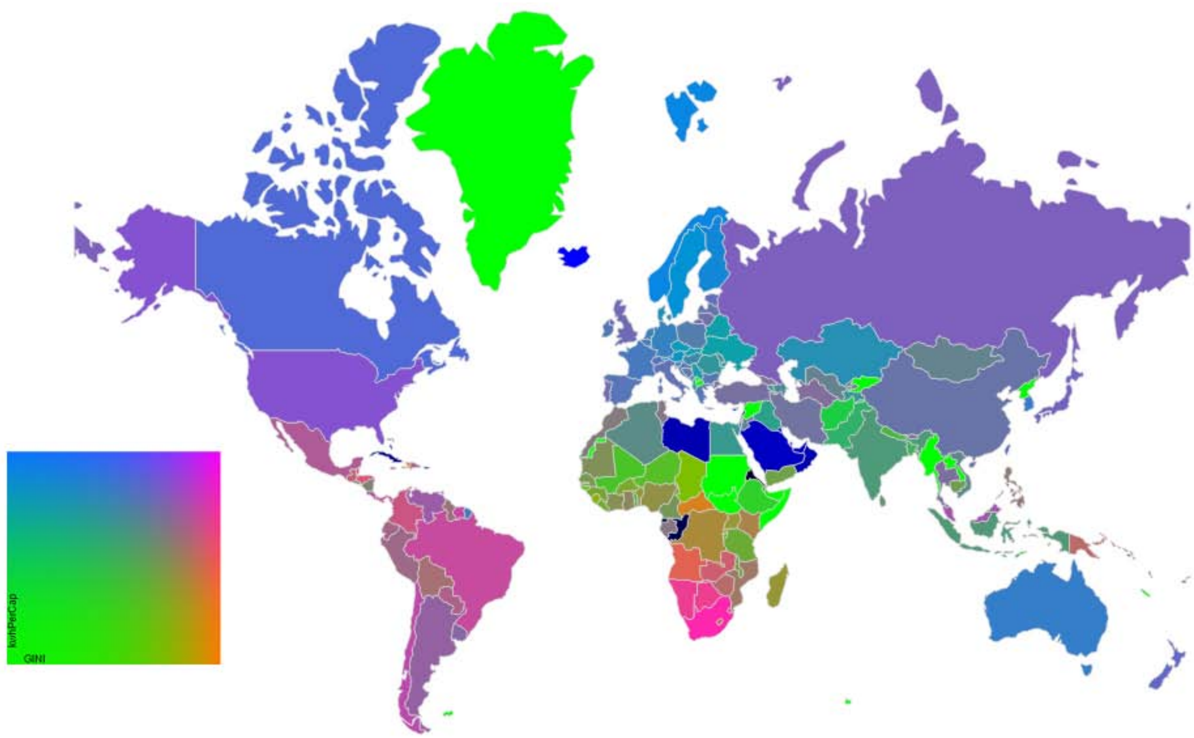


Figure 6-2: Contrasting Wealth Concentration with Electricity Use Per Capita

6.2.1 Model Generation

Files (objects) will be automatically created for use by the IEMM from the User Interface disaggregation process. Once the user is satisfied with the results of the disaggregation process, the electricity model will use the objects created from the Front End program to solve. These objects are passed to the Object Factory Manager. This component manages creation of the instances of the objects and their parallel solving.

6.2.2 Custom Reporting

Standard (canned tables) output reporting will be via CSV files at both the disaggregate level and at the WEPS+ 16 regions which would be produced from an output SQLite database. This SQLite database would also be available to the user for ad hoc queries. This information will also be able to be displayed in the GIS to allow the user to visually review the results. In addition, output electricity model data will be aggregated (summing and/or averaging as is necessary) to the 16 IEO regions that are needed for the rest of the WEPS+ and placed into the binary restart file (including electricity capacity, generation, consumption, and electricity prices).

7 Implementation Approach

This section describes the software packages to be used for the various facets of the model preparation. These packages must work together to complete the following tasks:

1. Accept user input from either a front-end GIS system (or a control file if running integrated with the rest of WEPS+)
2. Read in data sets and prepare data for use for both the LP models and simulation models based on user-specified parameters
3. Generate regional problem instances
4. Solve LPs and run simulations in parallel
5. Prepare solution results
6. Write solution results to a central database
7. Display results in the GIS system

7.1 Software Considerations

The IEMM will operate within the WEPS+ environment, so the computer languages and input/output structures need to be compatible with its Python shell. We plan on using Python being the primary language with likely calls to Python libraries in support of the simulation and the optimization. Because WEPS+ cycles in order to reach convergence among the supply and demand models, considerations of run time of the new IEMM will be incorporated into the model design.

The user interface will be constructed to facilitate the ability of the user to run alternative scenarios and sensitivity cases, update data, and modify the regional aggregation as desired. Scenario specification may involve modifying parameters for a single region, or for all regions at once. For example, alternative technology cost and performance assumptions may be used that address the uncertainty of technology costs globally or be specific to applications in certain regions.

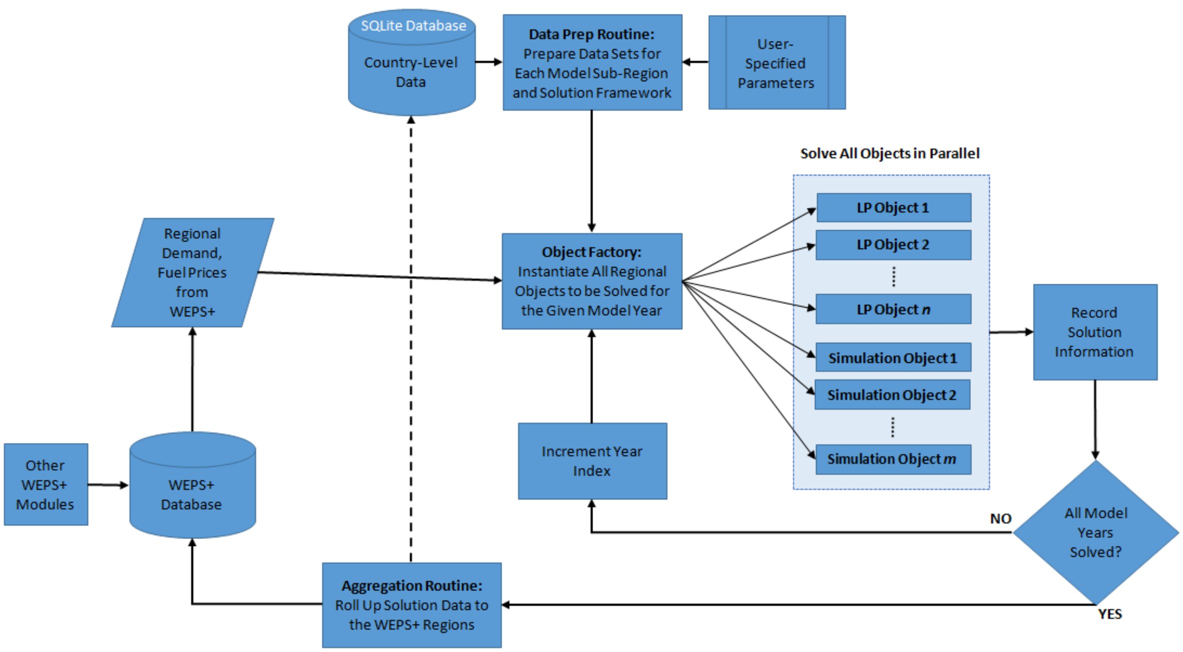
Data updates that are performed in support of EIA’s annual publication cycle should be relatively inexpensive to perform in terms of data acquisition and transformation into model inputs. In addition, all key data inputs will be organized in a logical and easy-to-access format. In order to have the flexibility to modify regional aggregation without making model code changes, most of the data will be provided to the model at the country level. The interface and accompanying model code will also use a flexible structure for technology specifications to make it easy to later add or subtract technologies.

Output reports are important for conveying results to analysts for diagnostic purposes and for summarizing into reports for outside audiences. While the standard reporting from WEPS+ will provide summary projections for the 16 regions, it will be important to have more detailed reports for diagnostic purposes. These will be constructed with the same flexibility to adjust as the sub-regional definitions are modified. The more detailed reports will also contain data on capacity, generation, and pricing that are not part of the standard reports.

7.2 Object-Oriented Programming in Python

We plan on leveraging Python's native object-oriented programming (OOP) capability to construct our modular solution approach. The regional breakdown of the IEMM lends itself nicely to this approach. Each region to be simulated or optimized would be a general object class, and then there will be 'simulation region' (either Detailed or Simple) and 'optimization' sub-classes which have their own associated data and solution techniques. Each regional object instance will be created in a data-driven way based on its data set and solution approach, and then solved in parallel. We will do this stepping forward in time to the end of the planning horizon. The general solution approach is summarized in the figure below.

Figure 7-1: Block Diagram of Model Components



7.3 Preparing the Input Data

Python has an extensive set of libraries dealing with the reading and writing of data in various formats, manipulating data, and connection to various databases (we are recommending SQLite in that it is already in use in the WEPS+ framework). We will use Python to read in and prepare the model input data for the simulation and LP model regional instances. The user-specified parameters will indicate the model regional aggregation level and what solution approach each region will employ, and the data sets will be prepared accordingly.

7.4 Auto-Generation of Model Objects

Once the data sets are prepared, Python will automatically instantiate regional model objects based on their solution approach and associated data sets. These model objects will then be passed to the simulation framework which will manage the object solutions over the model time horizon.

7.4.1 Simulation Objects

We plan to use Python as a simulation framework for executing the solution algorithm for the simulation region sub-class objections. Each simulation object with be instantiated with its own unique data set and then 'solved' using whatever solution algorithm is associated with this sub-class. The solution algorithms will also be written in Python.

7.4.2 Optimization Objects

The solution method for the optimization object sub-class will be needed to be an LP solver that can interface with Python. Fortunately, there are several current tools and packages for setting up and solving linear programming problems using open-source languages such as Python and applications libraries written in Python. These tools are able to call both open-source solvers as well as commercial solvers like CPLEX. These tools are mostly part of the Computational Infrastructure for Operations Research (COIN-OR) initiative which aims to spur the development of open-source software for the operations research community.[7]

Open-Source Optimization (Algebraic) Modeling Packages

The three main packages are described in Appendix A. The key to these packages is that they are open source and supported by a community of users. As just stated above, they can invoke a number of solvers both open source and proprietary (like CPLEX). As the selection of the language will have an impact on EIA's ability to maintain and enhance the resultant model with its own staff, we expect to work closely with EIA in determining the most suitable package.

Initially we are recommending the use of **Pyomo** (http://www.pyomo.org). However, the final selection will be made in consultation with EIA to assure it meets their standards and long-term objectives.

---

[7] http://www.coin-or.org/index.html

Pyomo can be used to define symbolic problems, create concrete problem instances, and solve these instances with standard solvers. Pyomo provides a capability that is commonly associated with algebraic modeling languages such as GAMS and AIMMS, but also takes advantage of Python's full-featured high-level programming language with a rich set of supporting libraries.

Open-Source Solvers

There are several open source LP solvers available that can interface with Python. Pyomo, in addition to being able to use proprietary solvers, is attractive because it can also make use of the most widely-used open-source one, including:

COIN-OR Branch-and-Cut (COIN-CBC) (https://projects.coin-or.org/Cbc)
A linear and mixed integer programming solver written in C++. It can be used as a callable library or as a stand-alone executable. It can be called through AMPL (natively), GAMS, MPL (through the CoinMP project), AIMMS (through the AIMMSlinks project), PuLP, CMPL, OpenSolver for Excel, or JuMP. Note that this free solver in included in any GAMS installation.

CoinMP (https://projects.coin-or.org/CoinMP) is an C-API library that supports most of the functionality of COIN-CBC. When compiled for Windows it generates a CoinMP.dll that can be readily used in other Windows C/C++ projects. When compiled for Unix it generates a CoinMP.so library that can be similarly used in other Unix C/C++ projects. The project includes a precompiled CoinMP.dll that is ready to be used as is in other Windows applications.

Interior Point Optimizer (IPOPT) (https://projects.coin-or.org/Ipopt)
IPOPT is a software package for large-scale linear and nonlinear optimization and is part of the COIN-OR Initiative. It also ships free with GAMS.

Again, while there may be some technical reason why we or EIA prefers one of these over the others, or would rather just use its existing solver licenses, this is something that will be worked out with EIA before proceeding.

7.5 Aggregation and Reporting

When the IEMM cycle is finished, there will be a series of routines to collect the sub-regional output and to roll it up to WEPS regions (summing, averaging, etc.). Further, the data will be associated back to the object so its linkage to the countries and WEPS regions can be displayed and reviewed.

8 Uncertainty and limitations

One of the huge uncertainties facing the development of an electricity capacity and generation forecast is the rate at which some region/countries will grow. Not only does China and India represent a huge challenge. Will their economy grow at 3-4% or 6-8% over the next twenty years in India where only a relatively small fraction of the population lives in cities today (32%-34%)? Migration to urban areas in India will fundamentally shift the need and type of electricity supply required to support it. In the example above, for Africa we used a simple (analyst judgment) GINI coefficient to skew the assumed level of demand towards cities where we believe greater incomes will lead to relatively larger electricity consumption.

Climate change and the world's response to it will pose a significant uncertainty for estimates of capacity and fuel consumption in the electric markets. What funding mechanisms will be put in place to aid the less developed and developing countries move to a more carbon neutral generation?

There is considerable turmoil in the oil and gas sectors that is having and will continue to have an impact on available supplies. The move to LNG will also change the power generation options, particularly in places like South Korea, Japan and China.

Each of these uncertainties will bring to the surface the limitations of any given electricity forecast. The only way to address them is to run a number of scenarios across the spectrum of the key drivers.

The design of the IEMM described in this paper will not only facilitate that analysis, but it will allow greater depth in understanding the impact as it looks into the sub-regional results.

9 Recommendations

The model design described in this document is based on a commitment to a responsive staged-design approach. This design process incorporates an evolutionary development schedule with delivery of each phase as early as possible, allowing EIA to provide feedback and suggest changes resulting in continuous improvements as the system is being built out. We will be responsible for the design and development working in close coordination with EIA. With this in mind, the next steps would include:

- Discussions with EIA regarding the design
    - Software selection (SQLite, Python, selected Python libraries)
    - Optimization Solver (Open Source or Proprietary)
- Development of a Process for Capturing Feedback/Change Request from EIA and Incorporation in Design
- Development of Schedule for Rapid Release of Key Prototype Components
    - Front-end Interface
        - Database Design and Implementation
        - User Interface Design and Implementation
    - Main IEMM Module Implementation
        - Main Routine
        - Object Factory Manager
        - Solution Subclasses
            1. Optimization model
            2. Detailed simulation model
            3. Simple simulation model
        - Back-end Aggregation and Reporting

The responsive design is implemented by swiftly producing a functional model to initiate a feedback loop with EIA and to identify changes that better suit one's needs. Figure 9-1 provides an overview of this process.

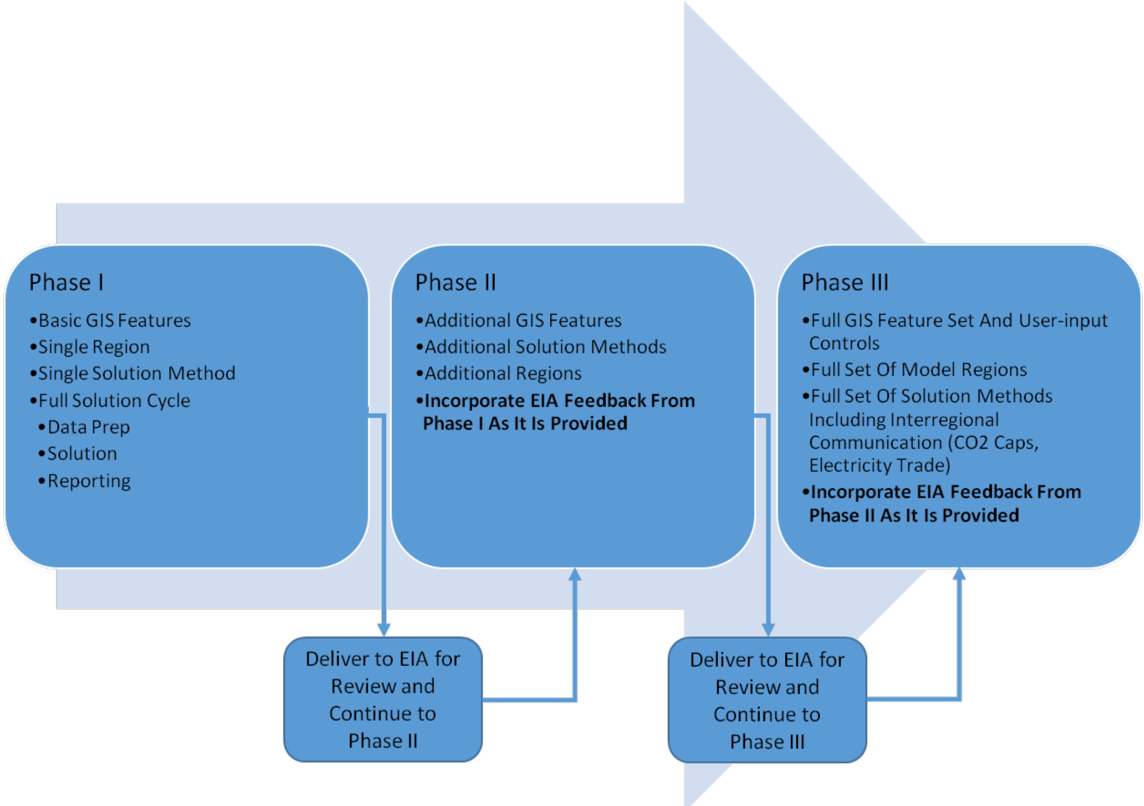


Figure 9-1: Overview of Responsive Design Process

9.1 Pre-Development Software Design Review

Prior to the initial software prototyping phase, we would solicit feedback from EIA modelers regarding our proposed model design as part of pre-phase initial design review process. This may involve suggestions regarding additional data sources, modifications to algorithms, addition of features, as well as LP solver selection criteria. A series of meetings which include detailed discussions to address specific design elements is our preferred path forward.

9.2 Phase I Prototype

The first phase of software development would involve building a stand-alone prototype in which all core design elements are constructed and tested. Specifically, a basic GIS feature set will be implemented, as will a full solution cycle capability (data preparation -> object generation -> solver execution -> reporting).

The prototype would focus on one of the 16 WEPS+ regions and would demonstrate the flexible design in creating appropriate sub-regions for the IEMM. An SQLite database would be constructed for the countries within that WEPS+ region that would be used by the Data Preparation Routine. The key elements performed by the routine will be developed, such as the IEMM regional definitions and associated data compilation (capacity initialization, demand shares, etc.) and load shape formation.

The Object Factory Manager would be developed to take information from the Data Preparation Routine and launch the various simulation and optimization models. For the Phase I prototype, the general structure would be built to allow the use of various types of solution approaches, although only one solution method will be fully implemented in this phase. In addition, the framework for iterating from year-to-year and aggregating back to the WEPS+ region would be constructed.

In Phase I, a prototype for only one of the solution methodologies will be created, most likely the Detailed Simulation approach. The Optimization and Simple Simulation models would be created in Phase II. As these are being constructed and tested, experiments can be made with some of the choices of dimensionality (such as load slices and segments). We can also test how the region selected for the LP approach performs using the simulation approach.

The last step is to construct the reporting back end in which the IEMM region results are aggregated into the WEPS+ region results and formatted output results are constructed.

After the Phase I prototype has been completed, it will be tested with a set of alternative scenario assumptions to test its robustness and response to a variety of market conditions.

At this point the prototype will be delivered and presented to EIA for review and comment. Work on Phase II will proceed while EIA digests the Phase I results and prepares a set of recommendations.

9.3 Phase II

In Phase II of the IEMM development, the feature set on the foundation built in Phase I will be expanded as will the number of model regions in the solution set. Recommendations from EIA, stemming from the Phase I prototype, will be incorporated at this time.

Additional front-end GIS and user-input features relating to the visualization of the input data and the selection of certain model parameters (such as the solution approach for each region or sub-region) will also be added during Phase II.

Additionally in Phase II, the set of solution options will be expanded to include the linear programming approach and an additional simulation approach. Feedback from EIA on the selection of LP software and solvers will be used to make sure they conform to EIA WEPS+ requirements.

Additional test scenarios will be added, at least some of which should include carbon emission taxes or penalty pricing. The purpose of these tests is only to verify the electricity model's response, rather than to assess the full economic response that would occur when using the full WEPS+ model.

As with Phase I, Phase II implementation will be delivered and presented to EIA for review and comment. Work on Phase III will commence while EIA digests the Phase II results and prepares a set of recommendations.

9.4 Phase III

Phase III involves the full model implementation. This phase will include additional GIS and user input features based on EIA feedback. It will also include the full WEPS+ model regional set and user-selected sub-regions. Furthermore, the Phase III implementation should include solution methods which allow for communication between regions as needed to model electricity trade or a multi-region carbon cap-and-trade policy.

At this stage, a major portion of the effort is data gathering to complete the dataset globally. In addition, a full benchmarking exercise will be needed to be performed on the EIA international energy balances for historic years.

Phase III will be delivered to EIA as the final prototype product for review.

Appendices

Appendix A: MODEL COMPARISONS

As EIA embarks on constructing a new international electricity model for WEPS+, it is useful to examine other global electricity models that are similar in the types of forecasts and policy questions they address.

A short description of five global integrated energy models is provided here along with their key attributes.[8] Four of the five models are simulation models, and the remaining model (MARKAL) is an optimization model. The focus is on the electricity and renewable components of these models which exchange price and quantity information with separate fuel supply and demand modules. All of the models use a bottom-up approach with technology options represented explicitly.

IEA World Energy Model (WEM). The WEM is a simulation model that covers energy supply, energy transformation and energy demand of the global energy system. It is the principal tool used to generate projections for the IEA World Energy Outlook scenarios, including global warming scenarios. Most of the model's data are obtained from IEA databases of energy and economic statistics.

TIMER (The IMAGE Energy Regional Model): TIMER is the global energy module of the IMAGE environmental assessment modeling system and is developed and maintained by the Netherlands Environmental Assessment Agency. TIMER is a technology-rich simulation model. The IMAGE results have played a key role in several global environmental studies by the IPCC and UNEP.

Prospective Outlook for Long-term Energy Systems (POLES): POLES is a global recursive dynamic simulation model of the global energy system that has been used by the European Commission and the World Energy Council to examine alternative energy and climate policy scenarios.

Global Change Assessment Model (GCAM): Developed by Pacific Northwest National Laboratory, GCAM is a dynamic-recursive market equilibrium model that has been used in a number of national and international modeling activities such as the Stanford Energy Modeling Forum, the U.S. Climate Change Technology Program, the U.S. Climate Change Science Program, and IPCC assessment reports.

[8] For a more complete description see "World Electricity Models: Final Report," prepared for EIA by OnLocation, Inc., August 8, 2014.

MARKet ALlocation Model (MARKAL)/TIMES: Originally developed at Brookhaven National Laboratory, MARKAL is now maintained and licensed by IEA Energy Technology Systems Analysis Program (ETSAP). The MARKAL-based models are technology explicit, dynamic partial equilibrium models of energy markets; equilibrium is obtained by maximizing the total surplus of consumers and suppliers using linear programming. The TIMES model is a later version developed by ETSAP that provides greater flexibility such as flexible time periods.

These models are compared based on the following five key features that are necessary in any electricity model and that would be expected to play a crucial role in a model's response to carbon prices and other alternative energy futures.

Technology Representation

All the models reviewed track existing capacity (including retirements) by technology and region and build new capacity as needed from a slate of new technologies that are competed based on costs, efficiencies and load characteristics. New technologies include a variety of fossil, nuclear and renewable technologies which may be affected by resource constraints imposed by the fuel supply modules. Differences exist between the models primarily in the number of technology types represented, and the level of detail used to characterize the carbon capture and storage (CCS), renewable and storage technologies which may be key players in a carbon price scenario. A biomass CCS option that produces "net negative" greenhouse gas emissions is available in three of the models (TIMER, GCAM and MARKAL) and is especially useful in a stringent carbon price scenario. Due to the intermittent nature of solar and wind, most of the models impose additional costs and/or limits reflecting the need for adequate reserve capacity.

Capacity Expansion, Retirements, and Retrofits

The projected need for new capacity (or generation) is driven primarily by growth in electricity demand in all of the models. Many of the simulation models decrement existing capacity using fixed retirement rates and add exogenous new builds if needed, then compare the resulting capacity to peak load plus a reserve margin to determine the need for new capacity. GCAM uses generation as the stock variable, with capacity reported as an output calculated using average capacity factors. Only the optimization model (TIMES) has a fully-integrated economic determination of building new capacity and replacing existing capacity, although GCAM and IEA WEM use economic indicators to either temporarily reduce generation or retire non-profitable plants. In most of the models that build to a capacity requirement, intermittent renewable capacity receives only a partial capacity credit that declines as its market share increases. Only some versions of MARKAL/TIMES and IEA WEM appear to have a CCS retrofit option.

Technology Choice for New Capacity

Two key facets required by any technology choice algorithm are diversity of technologies to reflect the range of loads, and sensitivity to fuel and environmental permit prices. The selection of new capacity and/or generation to meet increasing electricity demands is performed using essentially one of two different methodologies. The MARKAL/TIMES models use linear programming to minimize the cost of meeting electricity demand and to solve for the lowest cost mix of technologies, and perform an inter-temporal optimization across all projection years. The simulation models use logit functions to project market shares among technologies based on levelized costs (LCOEs). Several of the models (TIMER, POLES, and GCAM) also include parameters for non-cost factors that are used to calibrate the market shares. Levelized costs and shares are developed by load segment, where the number of mostly horizontal load segments varies from two (base and peak) in TIMER and IEA WEM, to seven in POLES. The exception is GCAM which operates in a single annual load segment.

Capacity Dispatch

Electricity generation and fuel consumption are key outputs of any electricity model, and allowing the generation mix of technologies to adjust to changes in fuel prices (and any associated carbon prices) is an important response in policy scenarios. The models we reviewed fall into three general categories of dispatch algorithms:

<u>Generation-Based Models</u>. GCAM is generation-based rather than tracking capacity and then performing a dispatch to determine generation. Generation from prior periods is brought forward in time after being decremented by retirements, and new generation is determined in the technology choice algorithms based on levelized costs. A ‘profit shutdown’ option can reduce a technology’s generation based on variable costs, thus giving this method some characteristics of a merit order dispatch.
<u>Merit Order Dispatch</u>**.** Most of the simulation models employ a merit order dispatch using variable costs to adjust the generation mix from both existing and new capacity in response to fuel price changes and carbon prices. TIMER and POLES dispatch intermittent renewable generation, CHP, and nuclear first, and the dispatch of fossil fuel plants is determined using variable costs and non-cost factors in a logistic equation in each load segment.
<u>Optimized Dispatch</u>. MARKAL uses linear programming to optimize the dispatch of plants and minimize total energy or electric system costs, subject to availability factors and other technology operating limitations. In a carbon scenario, this dispatch methodology will give priority to generation from low-carbon technologies.

Electricity Prices

Methodologies used to calculate end-use electricity prices fall into two basic categories, although some models take a hybrid approach. Average cost pricing is used in most of the simulation models, where electricity prices are based on average production costs (including capital) plus adders for T&D, taxes and, in some cases, other costs. In GCAM, the average price is based on the weighted average LCOE's for new plants. MARKAL optimizes across the energy system, and electricity prices are implicit in the shadow prices of the LP constraints. IEA WEM takes a hybrid approach and computes average marginal energy costs with an adder to ensure cost recovery for new plants.

Appendix B: OPEN SOURCE PACKAGES FOR OPTIMIZATION PROBLEMS

The key to these packages is that they are *open source* and supported by a community of users. As stated in the body of the report, they can invoke a number of solvers, both open source and proprietary (like CPLEX). As the selection of the language will have an impact on EIA's ability to maintain and enhance the resultant model with its own staff, we expect to work closely with EIA in determining the most suitable package.

Pyomo (http://www.pyomo.org). Pyomo can be used to define symbolic problems, create concrete problem instances, and solve these instances with standard solvers. Pyomo provides a capability that is commonly associated with algebraic modeling languages such as GAMS and AIMMS, but also takes advantage of Python's full-featured high-level programming language with a rich set of supporting libraries. Pyomo is supported by Sandia National Laboratory and is open-source. Pyomo supports a wide range of problem types, including, of course, linear programming.

Coliop|Coin Mathematical Programming Language (CMPL) (https://projects.coin-or.org/Cmpl)
CMPL is a mathematical programming language and a system for mathematical programming and optimization of linear optimization problems. CMPL executes the open-source COIN-OR Branch-and-Cut (COIN-CBC) solver by default, but can also use alternative solvers like GLPK, SCIP, Gurobi and CPLEX directly to solve the generated model instance. Since it is also possible to transform the mathematical problem into MPS, Free-MPS or OSiL files, alternative solvers can be used.

- The CMPL distribution contains Coliop, which is an IDE (Integrated Development Environment) for CMPL
- The CMPL package also contains PyCMPL and CMPLServer
- PyCMPL is the CMPL application programming interface (API) for Python. The main idea of this API is to define sets and parameters within the user application, to start and control the solving process, and to read the solution(s) into the application if the problem is feasible. All variables, objective functions and constraints are defined in CMPL.
- CMPLServer is an XML-RPC-based web service for distributed and grid optimization that can be used with CMPL and pyCMPL. It is reasonable to solve large models remotely on the CMPLServer that is installed on a high performance system. CMPL provides four XML-based file formats for the communication between a CMPLServer and its clients.

PuLP (https://github.com/coin-or/pulp) (docs at https://pythonhosted.org/PuLP)
PuLP is another open-source symbolic LP modeler written in Python. PuLP can generate MPS or LP files and call GLPK, COIN-CBC, CPLEX, or GUROBI to solve linear problem instances.

Appendix C: OPEN-SOURCE SOLVERS

There are several open source LP solvers available that can interface with Python. The most widely used ones seem to be:

COIN-OR Branch-and-Cut (COIN-CBC) (https://projects.coin-or.org/Cbc)
A linear and mixed integer programming solver written in C++. It can be used as a callable library or as a stand-alone executable. It can be called through AMPL (natively), GAMS, MPL (through the CoinMP project), AIMMS (through the AIMMSlinks project), PuLP, CMPL, OpenSolver for Excel, or JuMP. Note that this free solver in included in any GAMS installation.

CoinMP (https://projects.coin-or.org/CoinMP) is a C-API library that supports most of the functionality of COIN-CBC. When compiled for Windows it generates a CoinMP.dll that can be readily used in other Windows C/C++ projects. When compiled for Unix it generates a CoinMP.so library that can be similarly used in other Unix C/C++ projects. The project includes a precompiled CoinMP.dll that is ready to be used as is in other Windows applications.

Interior Point Optimizer (IPOPT) (https://projects.coin-or.org/Ipopt)
IPOPT is a software package for large-scale linear and nonlinear optimization and is part of the COIN-OR Initiative. It also ships free with GAMS.