International Coal Market Model
Component Design Report

April 18, 2017

Prepared for:
U.S. Energy Information Administration
Contract: DE-EI0003070

Prepared by:
José R. Benítez
Seth Schwartz
Energy Ventures Analysis, Inc.
1901 N. Moore Street, Suite 1200, Arlington, VA 22209-1706
(703) 276-8900
www.evainc.com

Table of Contents

Index of Figures

Index of Tables

Index of Equations

1 Executive Summary

This Component Design Report (CDR) contains recommendations from Energy Ventures Analysis (EVA) for the development of the International Coal Market Module (ICMM). The ICMM is a model within the World Energy Projection System (WEPS+) that will receive regional coal demands by regions, and, in response, it should provide other WEPS+ components with a forecast of coal prices.

EVA's proposed approach to represent international trade is based on the construction of regional supply-price curves. These curves are to be generated depending on the statistical production information available for each coal-producing country. Two types of supply-price curve approaches are proposed. The first type of representation, Class "A", is to be used for countries with available coal mine data. Class "A" curves should be constructed by representing the characteristics of each individual mine and extrapolating the data to potential future mines. The second supply-price curve type, Class "B", is to be used for countries that do not have enough data by individual mine. Class "B" curves should be constructed by statistical analysis of historical coal markets and their relation to macroeconomic indicators.

Coal supply and demand regions should be connected to each other using a transportation matrix. This matrix should represent the least-cost connection between the regions and will leverage the Ocean Freight Transportation developed by Hellerworx for EIA.

The supply-price curves and the transportation cost model are to be integrated into a Non-Linear Problem (NLP) that will solve for coal trade between regions by minimizing global system cost. The NLP approach is expected to be able to provide a stable solution in its price response because of its ability to solve continuous supply-price functions. Handling of model inputs, execution, and outputs will be organized using a Knowledge-Based Framework.

Finally, this report contains recommendations for ensuring uniformity in error handling and an organized reporting process. The proposed error-handling framework is based on a uniform categorization of errors that should lead to reduced model troubleshooting times. The proposed reporting process uses a categorized approach that separates reports between into categories: debug outputs, ICMM outputs, and WEPS+ common file outputs.

2 Introduction

Energy Ventures Analysis (EVA) has been tasked with developing a new analysis approach that can be used by the U.S. Energy Information Administration's (EIA) World Energy Projection System (WEPS+) to forecast international coal trade flows and resulting market prices. This approach needs to take into consideration not only the intrinsic properties of the product, but also the cost of connecting exporter with importer.

The modeling of international coal markets, at first glance, represents in many ways a typical forecasting problem where trade equilibrium of a commodity needs to be mathematically represented. This supply-demand problem quickly becomes more complex as a result of dissimilar coal properties and highly heterogeneous, or even missing, market information, and transportation infrastructure limits. Thus, coal cannot be seen as just one easily tradable commodity; any coal trade model needs to be able to recognize differences in heat content, coal constituents (i.e., sulfur, ash, moisture, etc.), and origin source.

To model international coal trade, this Component Design Report (CDR) proposes the creation of a model based on a least-cost global optimization of supply markets. The model would have as an input regional coal demand as calculated by WEPS+ energy use component models. Coal supply-price balance in a producing region would be represented by a statistically derived function. The regions would in turn be connected to each other using a least-cost transportation matrix. The end result of the model would be a price signal by use region that can be used for subsequent calculations.

This report also addresses issues associated with the development of a module. Recommendations are given in the areas of error handling, persistent in-memory variable management, and linking between Python and AIMMS.

3 Modeling of Regional Coal Supply-Price Response

3.1 Methodology Overview

Any modeling of international coal trade must, as a first step, consider that there is not one commodity but at a minimum four distinct types of coal. Therefore, the first step is to divide coal trade into four distinct markets according to major coal rank. These ranks are defined as metallurgical, bituminous, sub-bituminous, and lignite markets.

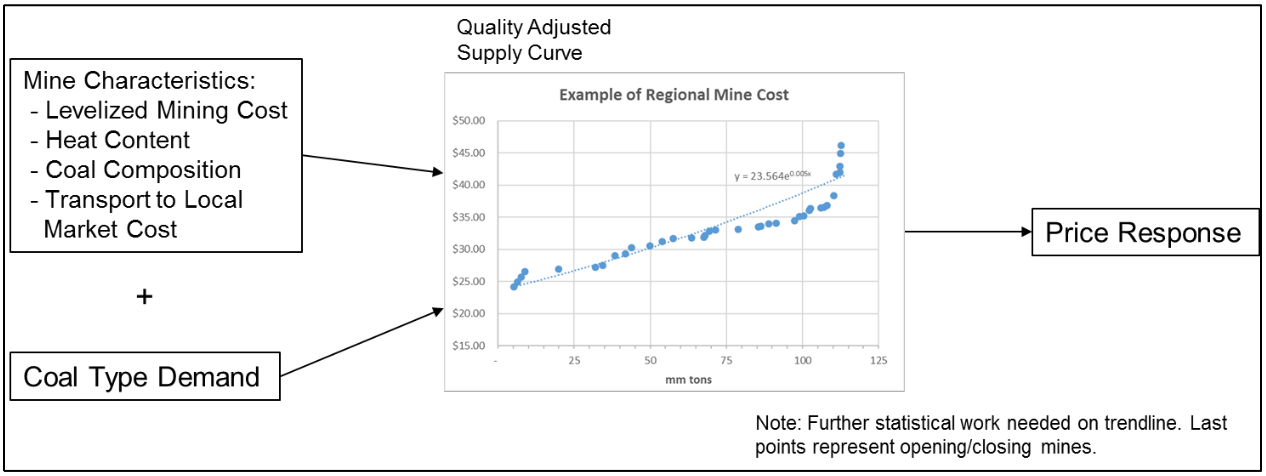
Furthermore, differences in coal composition would force the modeler to undertake a homogenization process in which coal prices from specific mines or regions are translated into market prices. EVA proposes this process to be undertaken by the use of Quality Adjusted Supply Curves. Such supply curves are to be created for each coal type, as needed, in each supply region.

In this report, we seek to address the modeling of the coal trade flow by first constructing supply-price curves to represent each coal supply region. These curves are constructed by either mine-cost analysis, or statistical analysis of supply-price relationships; these approaches are further explained in sections 3.3 and 3.4, respectively. Finally, supply-price representations are homogenized by using coal-quality adjustments as shown in section 3.2.

3.2 Quality Adjusted Supply Curve

The division of coal into markets by rank is generally not sufficient to compensate for the heterogeneities in coal composition properties. These differences in coal composition do have an effect in the final coal sale price. EVA proposes the use of Quality Adjusted supply curves to provide a framework to match the properties of coal supply and market quality demands. Coal prices are modified with factors to consider the following coal properties: (1) heat content, (2) sulfur, and (3) ash content. Figure 1 provides a representation of how such an adjustment would be integrated into a regional coal price.

Figure 1 - Example of Quality Adjusted Supply Curve



In the example above, the coal prices by mine are adjusted to reflect the pricing per ton on a market basis. This process is done by assigning adjusting coal prices based on the difference between mine and market coal specifications. The weight for the effect of each property difference is set by market analysis and modeler's judgment. The price modification in Equation 1 is applied to the price estimated for each mine to account for variations[1] in the heat, ash, and sulfur content, among other possible variables, from the base coal definition. The difference between market properties is defined as $Property_{Mine} - Property_{Market}$ and the Quality Adjustment Factor is set by analyst judgment.

Equation 1: Quality Price Adjustment Method

$$Price\ Modification = \frac{Property_{Mine} - Property_{Market}}{100.0 * Quality\ Adj.Factor}$$

The amount of coal produced in a mine is further adjusted depending on the implementation of coal-washing operations. If the coal product is washed, then an adjustment is made in the final quantity of coal produced.

All of these adjustment operations would maintain the physical coal-weight balance through the modeling efforts.

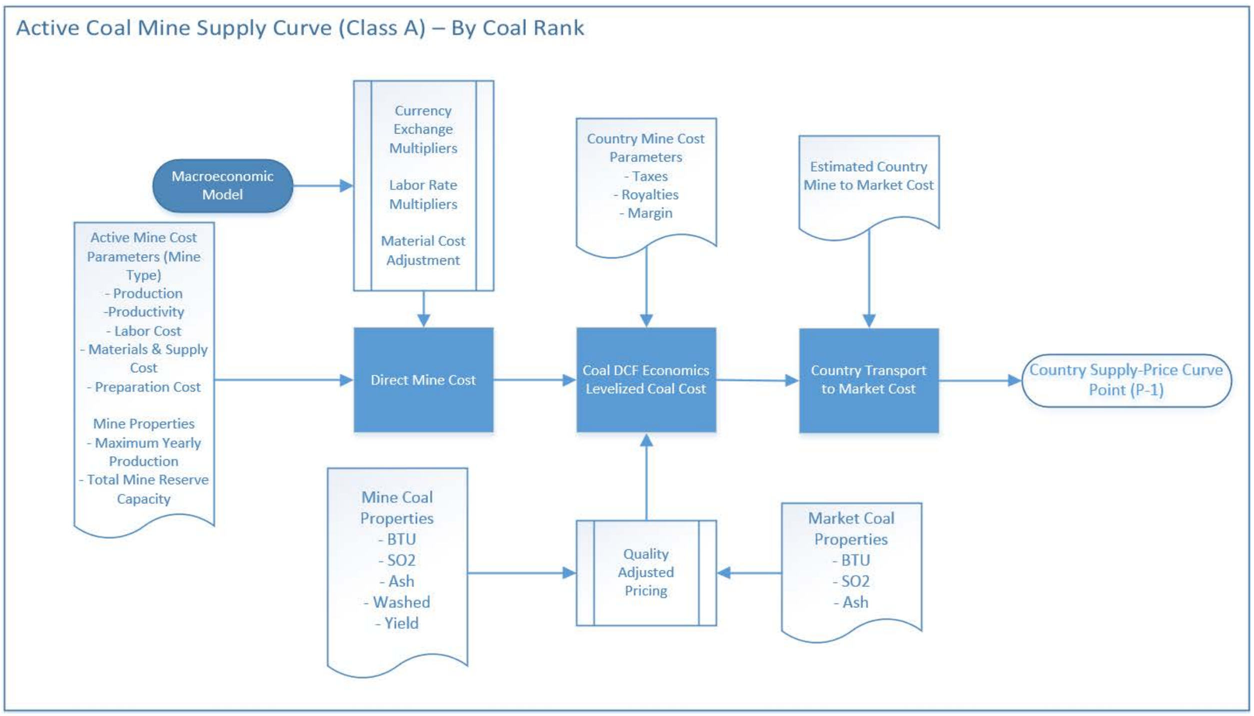
3.3 Class "A" Active Coal Mine and Future Potential Coal Mine Supply Points

This class has been created to represent coal prices in countries where detailed and accessible information about coal mine operations is available. The approach is a bottoms-up representation in which the individual behavior of active mines is represented based on an incremental mine cost modeling approach that allows the user to make assumptions regarding key trends in mine costs, macroeconomic variables, and country mine cost parameters as

[1] Further information on coal composition can be found at https://www.netl.doe.gov/energy-analyses/temp/QGESSDetailedCoalSpecifications_010112.pdf (Accessed April 14, 2017).

shown in Figure 2. This type of modeling permits the user to explicitly model different conditions by directly adjusting the appropriate variables. As shown in Figure 2, mine operational parameters are combined with macroeconomic variables to form the basis of the direct mine cost. The direct mine cost is then transformed into a levelized unit cost using regional parameters for taxes, royalties, and margin along with a transformation to market price by applying the previously described quality adjustments. Finally, the levelized mine coal price is transformed to a final market price by adding local transportation costs. The amount of coal produced by each mine is bound by both the maximum yearly production and total mine reserve capacity.

Figure 2 - Active Coal Mine Supply Curve (Class A)

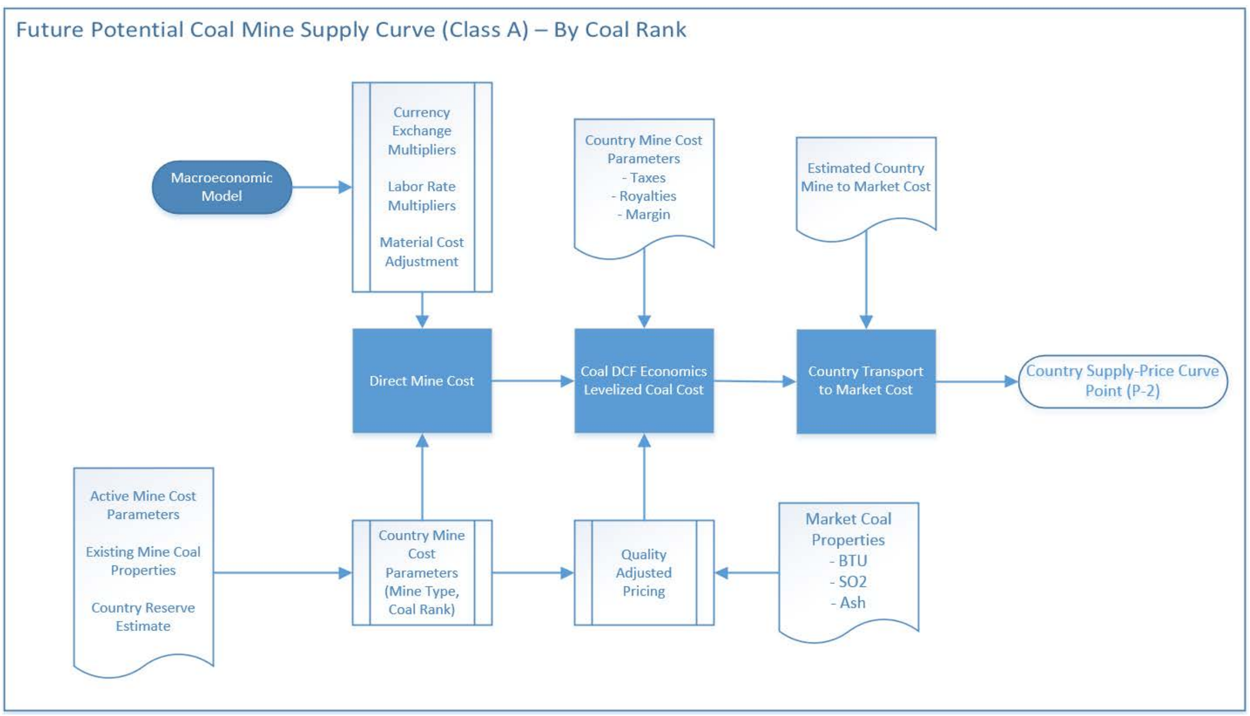


At this time, only the United States has been found to have sufficient public information to assemble a Class “A” representation. Another country that could potentially be represented in this manner is Indonesia, where worker productivity information is the only missing component. Australia also seems to have detailed information about its mines, but further research is needed to determine its suitability for Class “A” representation.

One advantage of the Class “A” representation is that the impact of technological changes can be readily assessed by directly varying the appropriate mine parameters. This assessment can enable the modeling and analysis of structural market shifts that might not otherwise be possible with a purely statistical representation. That assessment would then enable comparing and contrasting the results of Class “A” scenario modeling with other regions to determine the suitable selection of parameter for the other class representations.

Parameters for future potential coal mines would be determined through the aggregation of current mine parameters, as shown in Figure 3. That is, existing mine information would be used to generate the properties of future potential mines whose properties are not yet known because they don’t yet exist. Current mine parameters combined with detailed country resource assessments would provide the basis for the creation of future potential direct mine costs. Similar to existing mines, the direct-mine costs would be transformed into a levelized price by applying taxes, royalties, and margins. Quality adjustments and transportation costs will finally be added to convert the levelized mine cost into final market supply-price curve points. These new potential mines would be used to represent future mine development; the analyst would determine how many types of future potential mines are needed to represent a cross-section of future coal mine development.

Figure 3 - Future Potential Coal Mine Supply Curve (Class A)



Future changes in macroeconomic conditions, policies, and mine-development technologies affect both existing and future potential mine prices. The Class “A” approach permits the integration and analysis of these future changes. Provisions are to be made for these changes to be implementable via the use of time-dependent exogenous inputs for the model.

An average Mine-to-Market cost adder would be used to represent within-region transport costs (i.e., barge and rail shipping). This adder can be determined by specific analysis of the transportation links between individual mines and markets. Information on the cost of transportation of bulk goods via specific transportation mode is expected to be available for regions using the Class "A" supply curve representation.

3.4 Class "A" Supply Curve Creation

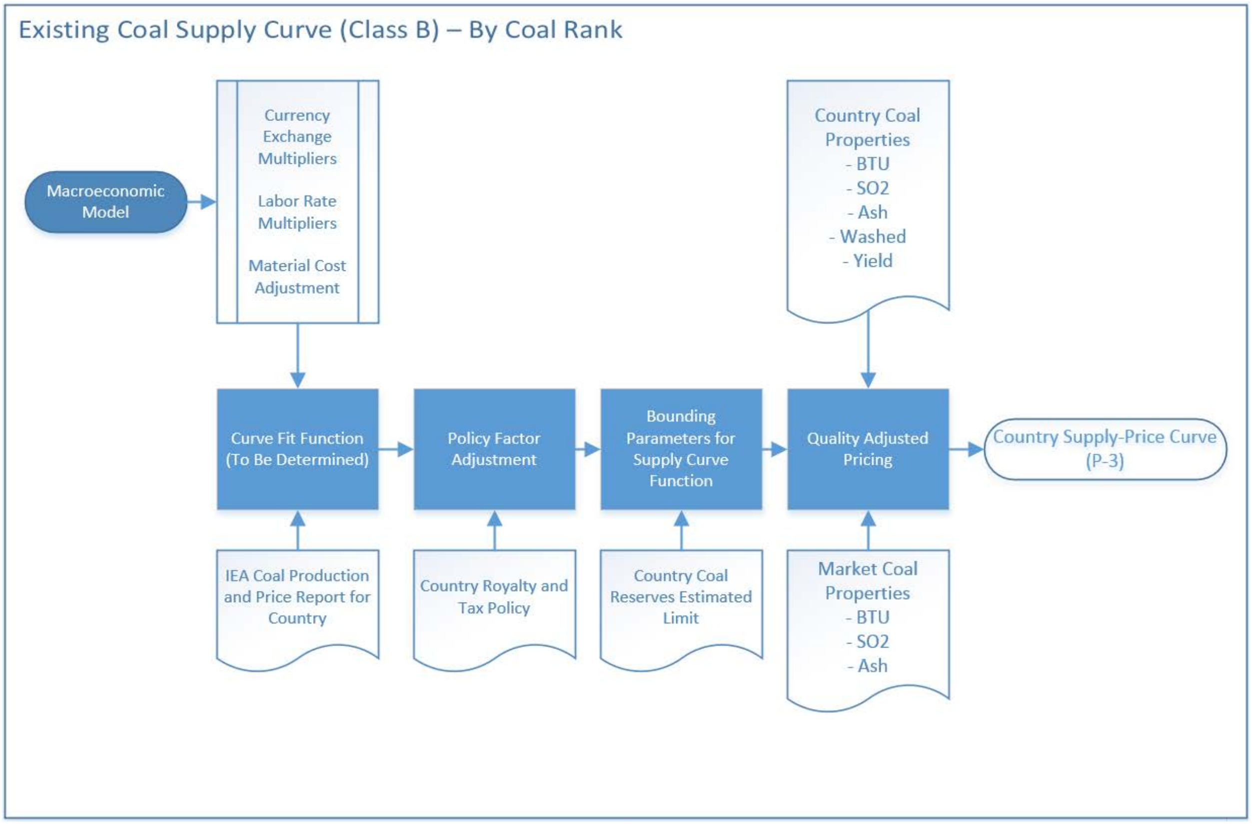
The Class "A" representation would provide specific price points for the supply of coal within a region. Those points then need to be transformed into a function that can be used to represent the market. The supply curve representing the market within the model would be assembled by applying a least-square curve fit over mine price points. The function used to represent the market would need to comply with convexity requirements as further explained in section 6.

3.5 Class "B" Coal Market Supply Curve

Class "B" representation is the modeling default for countries that do not have enough information to construct Class "A" representation; most countries fall under this approach.

As shown in Figure 4, the supply curve would be assembled using correlation coefficients optimized by least-square error minimization. The correlation coefficients are to be determined from analysis of variables identified as having a relevant impact on the price of coal. The supply curve function would still need to be determined through statistical data analysis, but if a NLP problem statement is used, then convexity requirements would apply. The likely functional form would be an exponential where the base and exponent are correlated to linear factors. The supply curves would be updated for each consecutive period to account for resource depletion and technological change.

Figure 4 - Existing Coal Supply Curve (Class B)



Priority should be given to correlation factors that are already part of the overall model (e.g., Macro Model Variables). The preference would be for endogenous variables since a forecast is already being established within the model.

Further study is needed to establish how technology change and coal depletion would be expected to modify the shape of the curve. Technical and economic resource assessments could aid in determining the trajectory of the supply-price curves by providing insight as to the tail end of the curves.

4 Overview of Coal Trade Information Sources by Major Importer/Exporter

4.1 Major Exporting Countries.

An important first step in the development of the International Coal Market Model (ICMM) is identifying the major producing countries in the coal trade market. This identification provides the basis for defining the coal supply regions within the ICMM. Tables 1 and 2 show coal supply and demand for major coal-exporting countries, by thermal and metallurgical coal types, for the year 2015 based on IEA's World Coal Outlook (International Energy Agency, 2016).

Table 1 World Thermal Coal Supply & Demand for Exporting Countries

World Thermal Coal Supply & Demand 2015					
million tonnes					
Primary Exporting Countries					
Region	**Country**	**Produced**	**Consumed**	**Exports**	**Imports**
Africa	South Africa	249	172	77	-
America, North	USA	691	642	31	9
America, South	Colombia	86	5	81	-
Asia	Australia	252	53	205	-
Asia	Indonesia	467	88	366	-
Europe	Russia	198	88	133	24
		1,943	**1,048**	**892**	**33**
Other Exporting Countries					
Region	**Country**	**Produced**	**Consumed**	**Exports**	**Imports**
America, North	Canada	25	27	2	4
Asia	Kazakhstan	85	61	24	-
Asia	Mongolia	6	4	-	-
Exporting Country Total		**2,058**	**1,140**	**919**	**36**

Table 2 World Metallurgical Coal Supply & Demand 2015 for Exporting Countries

World Metallurgical Coal Supply & Demand 2015					
million tonnes					
Region	**Country**	**Produced**	**Consumed**	**Exports**	**Imports**
Primary Exporting Countries					
America, North	Canada	26	3	28	4
America, North	USA	57	18	42	2
Asia	Australia	191	4	188	-
Europe	Russia	78	60	18	2
		352	**85**	**276**	**7**
Other Exporting Countries					
Region	**Country**	**Produced**	**Consumed**	**Exports**	**Imports**
Africa	Mozambique	4	-	3	-
Africa	South Africa	4	4	0	0
America, South	Colombia	5	3	1	-
Asia	Indonesia	3	-	3	-
Asia	Kazakhstan	16	15	1	-
Asia	Mongolia	13	-	13	-
Asia	New Zealand	1	0	1	-
Exporting Country Total		**398**	**107**	**299**	**8**

4.1.1 Indonesia

Indonesia was the largest exporter of thermal coal in 2015. Indonesia should be represented with a Class "B" curve. Enough information sources are available for the representation of the country as a single market (Petromindo.com, 2016). Indonesia could also be a good candidate for representation under a Class "A" framework if information sources about worker wages are found.

4.1.2 Australia

Enough information is available about Australia to construct a Class "B" curve representation. Further work is needed to establish if there are information sources to upgrade the country representation to a Class "A" framework. The country will need to be represented by two coal model regions to improve mine-to-market cost representation. The two regions will encompass the Queensland and New South Wales markets independently.

4.1.3 Russian Federation

The Russian federation should be represented by two regions to improve mine-to-market cost representation. Despite having two regions, each region would represent the same geographical extension with each region connected to an exclusive set of end locations. Each pseudo-region will represent the transportation from the mines to the markets on the east and west. This particular representation is necessary to reflect geographical transport considerations within the country.

4.1.4 Colombia

A preliminary review of International Energy Agency data indicates that enough information is available for the use of a Class “B” representation.

4.1.5 South Africa

A preliminary review of International Energy Agency data indicates that enough information is available for the use of a Class “B” representation.

4.1.6 United States

Sufficient information is available for Class “A” representation of the United States. The country should be split into Eastern and Western regions to improve the mine-to-market cost representation. Although there is current coal flow from west to east, the two regions would not be connected within the model.

4.1.7 Canada

A preliminary review of International Energy Agency data indicates that enough information is available for the use of a Class “B” representation.

4.1.8 Kazakhstan

Kazakhstan’s coal production would be represented within the Russian Federation supply region. The Russian Federation has one supply region that would expand to two pseudo-regions to represent the cost differences in mine-to-local market transportation. Kazakhstan coal production will be integrated with Russia’s coal production curve.

4.1.9 Mongolia

Mongolia’s coal production would be represented within China’s region, which is discussed in the next sub-section.

4.1.10 New Zealand

A preliminary review of International Energy Agency data indicates that enough information is available for the use of a Class “B” representation.

4.2 Major Importing Countries

While identification of major exporting coal countries is the first step towards defining coal supply regions, representing importing countries with major domestic coal production is equally important. The importance stems from the necessary balance between in-region production and imports to satisfy local demand. Local coal supply curves will be developed for these regions to ensure that proper trade balances are maintained. This representation of local production will enable an import country to become an export country when local production exceeds local demand. Major importing countries with significant domestic production would also form part of the coal regions within the ICMM. Tables 3 and 4 show coal supply and demand for major coal importing countries, by thermal and metallurgical coal types, for the year 2015 based on IEA's World Coal Outlook.

Table 3 World Thermal Coal Supply & Demand for Major Importing Countries

World Thermal Coal Supply & Demand 2015					
million tonnes					
Importing Countries with Large Domestic Production					
Region	**Country**	**Produced**	**Consumed**	**Exports**	**Imports**
America, North	Mexico	12	18	-	6
Asia	China	2,916	3,094	4	156
Asia	India	594	763	1	171
Asia	Vietnam	37	42	2	7
Europe	Poland	60	58	7	6
Europe	Ukraine	32	31	1	5
		3,650	**4,007**	**14**	**350**
Other Importing Regions					
Region	**Country**	**Produced**	**Consumed**	**Exports**	**Imports**
Africa		9	6	-	7
America, South		9	31	2	24
Asia		17	406	-	387
Europe		28	209	3	179
Mediterranean		-	19	-	19
Importing Country Total		**3,713**	**4,678**	**20**	**959**

Table 4 World Metallurgical Coal Supply & Demand for Major Importing Countries

World Metallurgical Coal Supply & Demand 2015					
million tonnes					
Importing Countries with Large Domestic Production					
Region	**Country**	**Produced**	**Consumed**	**Exports**	**Imports**
America, North	Mexico	3	5	-	2
Asia	China	611	658	1	48
Asia	India	55	106	-	51
Europe	Czech	4	4	2	2
Europe	Poland	13	13	2	3
Europe	Ukraine	6	15	0	10
		691	**801**	**6**	**115**
Other Importing Regions					
Region	**Country**	**Produced**	**Consumed**	**Exports**	**Imports**
Africa		0	1	-	1
America, South		-	12	-	12
Asia		1	97	-	97
Europe		5	50	-	43
Importing Country Total		**698**	**961**	**6**	**268**

4.2.1 China

China has significant coal consumption and significant domestic production. The balance between local production, imports, and demand has historically had major impacts in international coal markets because of China's market size. One significant challenge with the representation of China is the low quality of information available about its coal production capabilities and internal price points. Further work would need to be undertaken to generate enough data for the construction of a Class B supply-price curve.

4.2.2 India

India's demand has been rapidly growing over the past couple of years. Most of the country's coal is produced by a state-owned entity, while only a small portion of production can be attributed to private entities. India has never been a net exporter of coal, and it is not expected to be in the foreseeable future due to the need for local production to be subsidized by government. To correctly forecast India's interaction with the world coal market, it coal export tariffs that would strip Government subsidies from any potential exports would need to be implemented; otherwise the model might incorrectly determine coal exports to other regions from India. Due to state-control of coal production entities, further work to be necessary to generate enough data for the construction of a Class "B" supply-price curve.

4.2.3 Poland

Enough data is available from IEA to construct a Class “B” supply curve for Poland. Further available information about individual coal mining companies can be used to cross-check IEA data.

4.2.4 Ukraine

Coal production information for Ukraine would need to be extracted from reports from publicly traded companies that operate in the country. Enough information from these reports is available to construct Class “B” supply curves. Coal production information from the country might need to be processed by analysts due to high volatility in the data as a result of current conflicts.

4.3 Other Coal Regions

Regions not identified as major exporters or major importers with significant domestic coal production would be aggregated by WEPS+ representative region. These regions would not contain supply-price curves, and their aggregated demand, minus any already modeled ICMM region, would be represented as a coal-use-only region.

5 Inter-Regional Coal Transportation Cost

5.1 Overview of Transportation Costs

Transportation of coal between regions is generally carried out by rail, barge, or oceangoing vessels, with transportation by oceangoing vessels being the preferred method. Coal supply regions should be connected using the seaborne transportation cost model developed for EIA by Hellerworx (Hellerworx, 2016). This model should be sufficient to represent the major coal trade routes between regions.

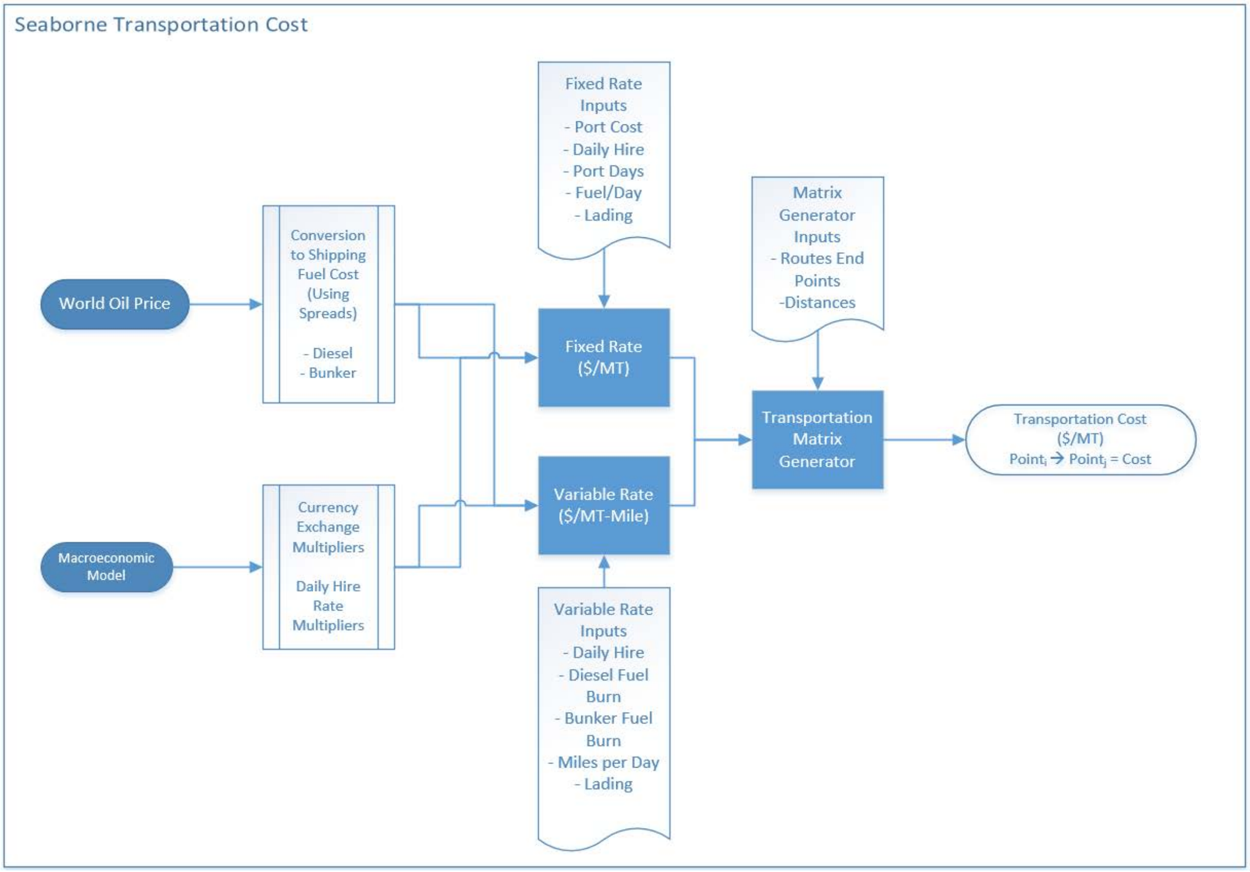
The port capacity should be explicitly represented within the model to correctly represent current and future coal trade chokepoints. Along with port capacity constraints, the model should also have the option to expand capacity by a set percentage when presented with a binding port capacity constraint. This allowance for expansion of port capacity will need to be tailored for each port to represent local conditions.

Future enhancement of the model could include the explicit representation of intra-regional coal trade using rail and barge. This enhancement is likely to be needed for a limited set of regions.

5.2 Seaborne Transportation Costs

Seaborne transportation costs are based on the proposed Hellerworx Ocean Freight Model developed for EIA (Hellerworx, 2016). The model considers transportation of coal in two distinct type of vessels: Panamax and Cape-size vessels. The transportation cost model considers as an input the price of distillate oil for port operations and residual oil for ocean-fairing operations. The cost per ton of coal per mile transported is higher for a Panamax vessel, but that vessel type can take advantage of shorter travel routes through the Suez and Panama canals. Figure 5 prepared by EVA shows how the Hellerworx Ocean Freight Model might be integrated into the ICMM.

Figure 5 - Seaborne Transportation Cost



It is important to note that the Panama Canal opened a new and bigger route on June 26, 2016. This new route enables the crossing of the Panama Canal by Post-Panamax size vessels. EVA has already seen an increase in Colombian coal exports now that larger sized vessels can pass through the Panama Canal. Further work is needed to integrate this new vessel size option into the Hellerworx Ocean Freight Model.

5.3 Rail and River Barge Transportation Costs

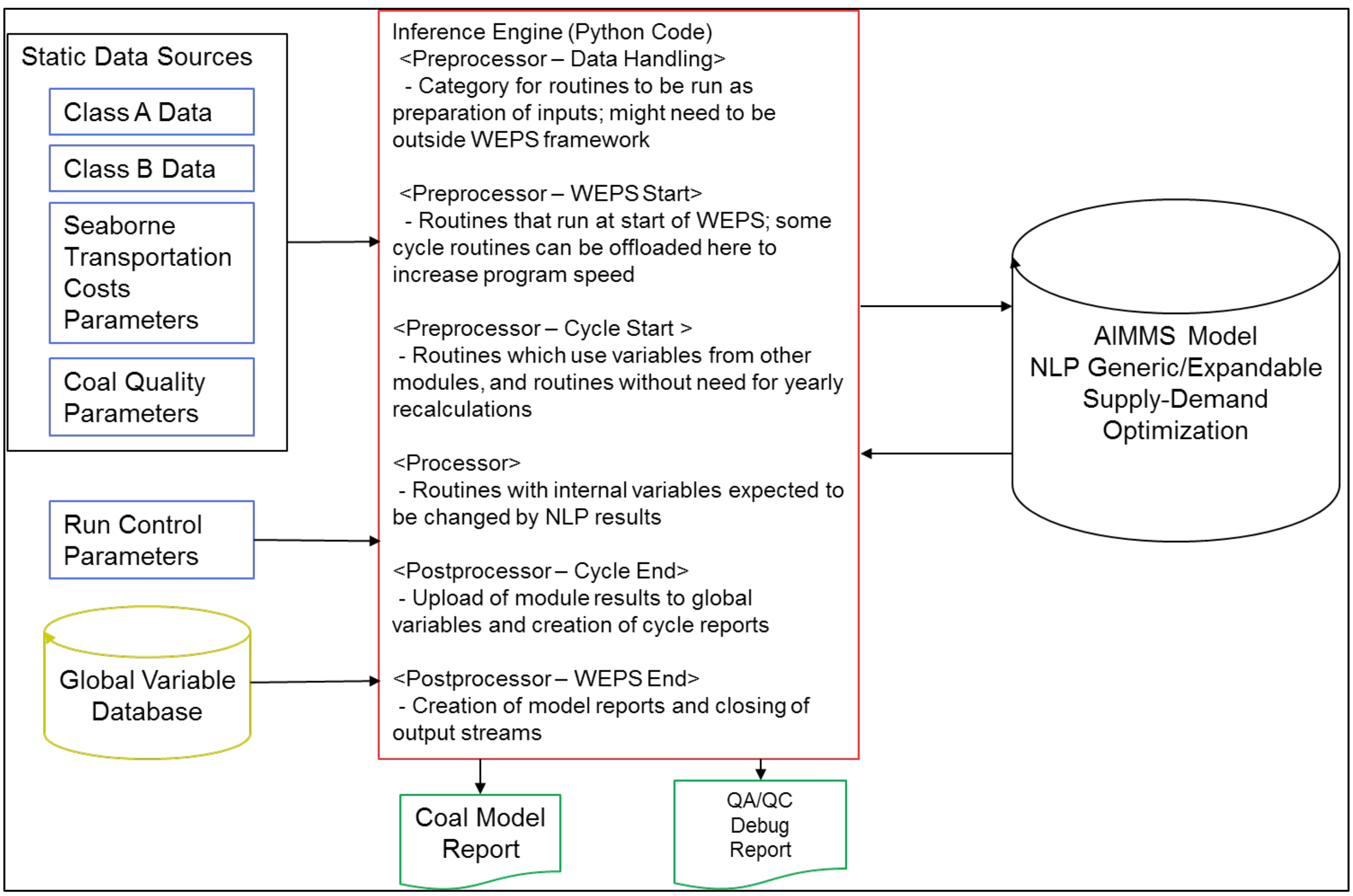
Rail and barge transportation costs may need to be considered in the future a higher degree of model resolution. This type of transportation option might be needed if supply and demand regions are further divided to increase model resolution.

6 Configuration of Main Trade Flow Representation

6.1 Knowledge-Based Model Framework

The overall structure of the International Coal Market Model (ICMM) will be based on the Knowledge-Based Model Framework. Under this framework routines, to the maximum extent possible, for the transformation of raw data sources would be stored within the ICMM; this in turn minimizes ad-hoc processing of information and standardizes the analysis of data. Routines for processing information have been divided into categories; these categories would house processes which have common starting/stop times during model execution. Figure 6 shows the organization of the ICMM. It is structured in accordance with Knowledge Based Modeling principles.

Figure 6 - Overall Model Framework



Static data sources comprise the main input to the model; these inputs are considered to be the raw input from primary data sources and data vendors. Static data information sources are meant to be kept as raw as possible to minimize the need for ad-hoc analysis. Not all data sources are amenable to this treatment; some data sources will need direct human analysis and manipulation before its use with the ICMM. The following is a categorized list of potential information sources ranked with the level of analyst interaction. An example of the information sources, and how they are handled within the ICMM, is given to illustrate the implementation of the Knowledge Based Framework.

- Fully Automated Primary Data – This type of information should be able to be fully automated and processed by the ICMM. Some formats will require special interface considerations due to the need for parsers and interpreters. Code for reading and processing this type of information will reside within the Inference Engine.
  - IEA International Coal Trade Data – This dataset is already available in a computer-readable format. It might require the use of an Application Programing Interface (API) for accessing the database format in which information is distributed[2].
  - Bloomberg Terminal – Some of the information required for the ICMM will need to be obtained from information vendors. As an example, historical performance of coal markets not covered by IEA and coal transport cost information can be accessed via the Bloomberg Terminal. Third party information providers, such as Bloomberg, make available their data via the use of an API[3]. A wrapper might need to be constructed for information sources that do not provide their API in a Python code format.
- Semi-Automated Primary Data – This type of information would need some level of human intervention due to the use of incompatible formats or lack of an offering of the information product in a machine-readable format. The amount of human intervention should be limited to data transfer only. It is envisioned for this type of information to be provided via the use of properly labeled spreadsheets that the ICMM inference engine would be able to process.
  - No information sources have been identified as falling in this category, but previously mentioned fully automated data sources might need to be handled this way if intractable computer integration or licensing problems arise.
- Analyst-Driven Data – This type of information will need substantial human involvement and might require the use of analyst judgment. This type of information should be used as a last resort and only when other types of information are not available. Use of this type of information is not strictly compliant with the Knowledge Based Modeling Framework; nonetheless, it might be needed where the use of a human analyst is unavoidable. Efforts should be made for the Analyst Driven Data to require minimal

[2] http://data.iea.org/payment/WDSTechnicalRequirements.aspx (Accessed April 14, 2017).
[3] https://www.bloomberg.com/professional/support/api-library/ (Accessed April 14, 2017).

updating. Finally, standardized data input spreadsheets should be used to minimize data entry errors.

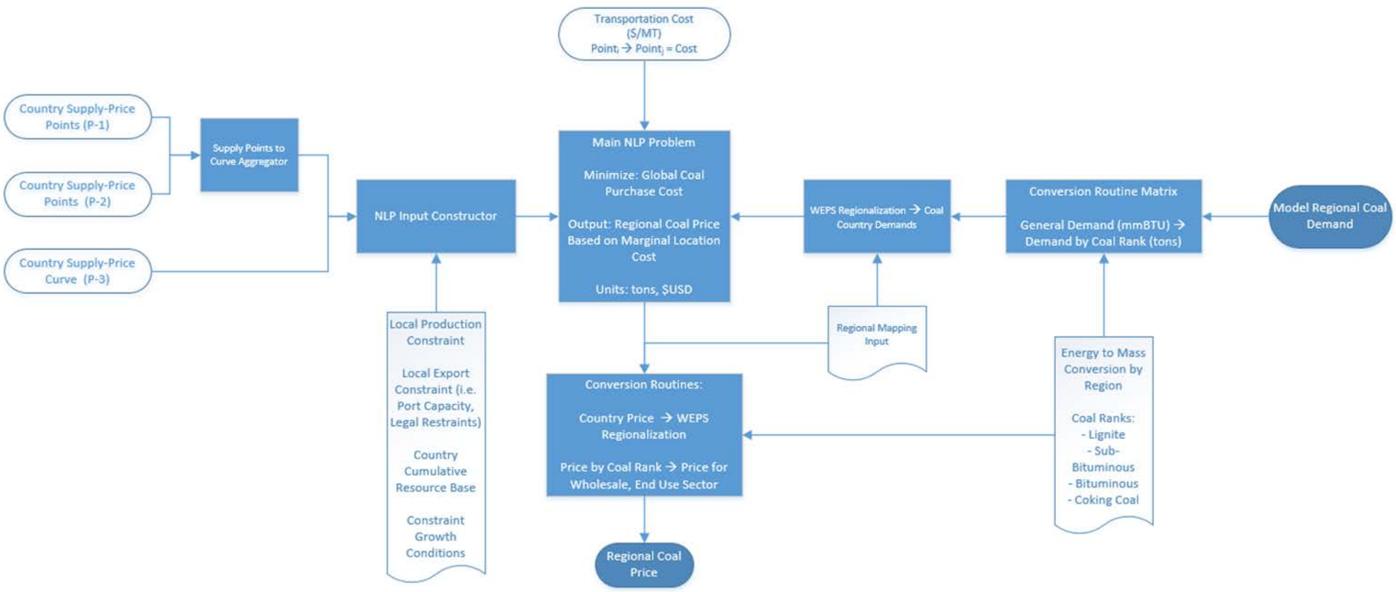
- Indonesia Coal Information – Information needed to describe coal mines in Indonesia is available in paper format and will need to be processed by a human analyst to determine the coal properties to be used for that country within the ICMM. Once this information is processed for coal mine characteristics, updates would only be needed on an ad-hoc basis as recently closed and opened mines accumulate.
- Run-Control Parameters – This is a set of controls that would need to be set by the analyst to instruct the ICMM on the type of analysis that would be needed. Not strictly a data source, but it is nonetheless an external information source to be used on the control of the ICMM run.

Standing between the information sources and the ICMM optimization engine in Figure 6 is the Inference Engine. The function of the Inference Engine is to process the data sources into a usable format that can be used by the optimization routines to determine coal trade flows. The inference engine also processes outputs from other models within WEPS+, like coal demand and macroeconomic variables, to be used as inputs within the coal trade model. Finally, the inference engine reads the ICMM optimization solution and prepares the reports associated with the model run. Section 6.2 provides an in-depth discussion on how the inference engine works and section 6.3 provides an in-depth discussion of the optimization problem formulation that forms the Supply-Demand optimization model.

6.2 Modeling Overview

The different components of the International Coal Market Model (ICMM), as shown in Figure 7 and described in the previous sections, would be connected to achieve a representation of the International Coal Markets. This representation provides a generic framework, independent of coal rank, regionality, and forecast year, which would be housed within the AIMMS component of the ICMM.

Figure 7 - Overall Representation of Coal Markets



The AIMMS component of the model would start its operations by receiving Class “A” and “B” representations of the supply-price curves by ICMM coal region. Class “A” mine points will be aggregated and converted into a continuous, differentiable, and convex supply-price equation via the use of a least-squares curve fit. That is, the individual mine supply points will be aggregated via the use of a mathematical function to represent a continuous supply-price function. These curves, combined with the identified constraints will be used to assemble the main NLP optimization model. It is important to note that in the formulation of the model for subsequent years, the supply curves would be modified to reflect technological changes and recognize cumulative coal depletions. Class “A” supply curves could be changed by direct modification of the mine properties and costs; the shape of the representative curve could be adjusted. Class “B” supply curves, due to their nature, can only represent future conditions via exogenous changes to curve magnitude and shape. Implementation of future changes will depend on the analysts’ views of potential future conditions.

The ICMM is expected to receive the coal demand aggregated at the WEPS+ region level. This level of aggregation will likely cause the demand for several ICMM regions to be lumped into a single WEPS+ region. To resolve individual ICMM demands, the model will need to apportion the WEPS+ demand into ICMM regions. To do this, the model should as a default contain a down conversion routine. This routine is to use a static predefined matrix of WEPS+ to ICMM demand allocation factors. Such factors will be pre-calculated by looking at the historical demand share for the WEPS+ and ICMM regions.

The transportation cost coal matrix, previously calculated by external to AIMMS routines, would be added to the NLP formulation to represent the cost of moving coal across ICMM region. Constraints associated with transport would apply across all coal ranks because transportation is agnostic to the type of coal being transported. Applying such a constraint would necessarily involve the linking[4] of coal rank representations within the model. Appropriate levers would be implemented to de-link the coal models by rank in case the user experiences prolong model solve times. When the de-link lever is activated, the model would allocate the overall coal transportation constraint based on historical coal rank export ratios.

Routines for the reporting and aggregation of coal trade flows and prices would run after the solution of the coal trade flow model is completed for all forecasting years. These routines would translate the model solution from ICMM regions to WEPS+ regions. The routines would work by allocating prices based on the share of supply by each ICMM region into the total coal supply for the entire WEPS+ region. The allocation for prices will be done as shown in Equation 2, where the price in a WEPS+ region composed of several ICMM regions will be weighted according to the coal supply generated by each ICMM region. In case of multiple WEPS+ regions being covered by a single ICMM region, the coal price will be set for all regions by the price at the ICMM level[5].

Equation 2: Up-Conversion of Outputs from ICMM to WEPS Regions

$$Price_{WEPS\ Region} = \sum \frac{\left(Price_{ICMM\ Region} * Supply_{ICMM\ Region}\right)}{Total\ Supply_{WEPS\ Region}}$$

[4] Generally, each coal rank is treated as an independent component within the model; that is, the supply of a particular coal rank within a region is independent of other coal ranks. The transportation of coal presents a particular problem to this separation of coal ranks due to the ability of transportation equipment to be agnostic as to the type of coal being transported. The sharing of transportation by different coal ranks will require the model to solve all the equations for all coal ranks simultaneously instead of individually. This means that the model will have to determine which coal gets transported when there is not enough port capacity for all. One potential approach to keep problems separated by coal rank is to apportion port capacity based on historical usage. This approach can minimize the computational solve time for the problem, but it can introduce errors if there is a significant shift in coal exports for a region.

[5] Although it is not expected for an ICMM region to cover multiple WEPS+ regions, this logic will be needed within the ICMM model to comply with the Knowledge Based Framework.

Finally, results for coal flow and prices by ICMM regions would be created for other models, which might require more detail for coal prices and quantities. More information about this output is described in Section 7 of this report.

6.3 Optimization Problem Formulation

EVA has developed two approaches that can solve for the optimal coal trade balance between regions that produce and consume coal. The first, and preferred approach, is via the setup and optimization of non-linear functions to model the supply-price characteristics of a particular supply region. These supply regions are then connected to other regions via a transportation cost adder matrix.

The second approach to solving the trade flow problem involves the use of regional coal supply bins as part of a linear problem (LP). Testing of the LP approach revealed that the solution was sensitive to the appropriate creation of the price bins. Although the use of price bins in combination with an LP problem has the advantage of offering unlimited supply price curves, this in turn comes with observed model instability due to *knife-edge* effects.

6.3.1 Non-Linear Problem Approach

The Non-Linear Problem (NLP) approach entails the representation of the global cost to produce and deliver coal through an objective function shown in Equation 3. The result of the objective function is then minimized while the problem constraints are met. The non-linear component of the problem stems from the use of a non-linear relationship to represent the price of coal in a region given a certain production level. The equation shows the relationship between produced coal and prices by the use of an exponential relationship. The objective function to be minimized is:

Equation 3: NLP Objective Function

$$\text{Cost} = \sum_{reg=1}^{Regions} [\![Quantity(reg_{from}, reg_{to}) * \left(Price(reg_{from})\right]\!] + Transport(reg_{from}, reg_{to})\Big)$$

The term *Quantity* represents the amount of coal that would be produced by a region to be used within the region or as an export to other regions. The market price of coal supplied by a particular region is shown in Equation 4; where *reg* refers to region under consideration, and Price(Q_0) represents the coal price at zero supply.

Equation 4: Definition of Supply-Price Curve based on Exponential Elasticities

$$\text{Price}(\text{reg}_{\text{to}}) = Price(Q_0) * e^{elasticity * \sum_{reg=1}^{Regions} Quantity(reg_{from}, reg_{to})}$$

The Supply-Price curve function needs to be based on a continuous, differentiable, and convex equation to ensure that the objective function remains within the realm of Convex Optimization. This is important because a Convex Optimization problem is solvable in a reasonable amount of time and resources (Boyd & Vandenberghe, 2004). The following is a partial list of convex equations that can be used to represent the Supply-Price relationship (Boyd & Vandenberghe, 2004):

- Exponential -> e^(ax)
- Powers - > x^(a), but a >= 1 or a <= 0
- Powers of Absolute Value -> |x|^(a), but a >= 1
- Logarithm -> log(x)
- Negative Entropy -> x log(x)
- A function in which convexity can be checked by f''(X) >= 0

Any addition to the problem formulation needs to be carefully checked for convexity impact. For example, adding a demand response function like f(x) = e^(-c *x) to address demand sensitivity to coal pricess will add local optimum points.

Coal supply in the objective function is connected to other regions via the use of the *Transport* parameter. This parameter represents the result of the least-cost option to transport coal from a producing region to a consumption region.

The amount of coal that must be met for a region is introduced as a constraint to the problem. In this case, the regional demand for coal needs to be satisfied by a combination of local production and imports, as shown in Equation 5. Counterintuitively, the amount of coal that is being exported to other regions does not need to be subtracted from the equation below. This amount of coal being exported is already considered by the objective function and the demand parameter for the receiving region.

Equation 5: NLP Local Demand Satisfaction Constraint

$$\mathrm{Demand(reg)} \leq Demand_{local}(reg) + Demand_{Import}(reg)$$

Finally, the production of coal for a particular supply region would need to be constrained. This cap represents the maximum amount of production that a region could sustain in a given year, considering the local infrastructure to transport coal (Equation 6). A binding constraint in a given year should represent a need to expand the infrastructure network. Such an expansion will be allowed for the following model year subject to a maximum infrastructure growth limit. A high-price slack variable will be added to allow the model to expand the in-year to prevent model infeasibility solutions. Use of the slack quantity variable by the model should trigger a warning event.

Equation 6: NLP Maximum Regional Supply Constraint

$$\text{Production}_{\max}(\text{reg}) \geq \sum_{reg=1}^{Regions} Quantity(reg_{from}, reg_{to}) - (Price_{Slack}(reg) * Quantity_{Slack}(Reg))$$

Further constraints could be introduced to segregate production and transportation constraints by each region. A review of coal trade information reveals that the implementation of a single production constraint should be sufficient to control the model.

6.3.2 Linear Problem Approach

The Linear Problem (LP) approach is similar to the previously explained NLP approach. The LP has an objective function that calculates the total cost for the supply and transportation of coal to meet the coal demand in all regions. The difference in approaches stems from the use of price bins, instead of convex functions, to represent the relationship between coal supply and prices. These bins are structured to represent a discrete amount of coal that would be available at a certain price. One advantage of using price bins is the ability to represent by proxy any supply and price function without the convexity constraints. A major disadvantage is the inherent resolution problem introduced by the use of bins; the creation of bins become an art by itself. The use of bins also potentially creates the *knife-edge* effect that is problematic for LP models.

The objective function stated in Equation 7 represents quantity and prices for each region but also by price bin.

Equation 7: LP Objective Function

$$\text{Cost} = \sum_{reg,\ pbin}^{Regions, Bins} Quantity\big(reg_{from}, reg_{to}, pbin\big) * [Price(reg_{from}, pbin) + Transport(reg_{from}, reg_{to}, pbin)]$$

Demand and production constraints for the LP approach are the same as the NLP approach. Equation 8 shows the demand satisfaction constraint, which is similar to the constraint used in the NLP approach. Additionally, a constraint not previously required by the NLP approach (Equation 9) is included, which requires that the local regional demand to be met first by local production and imports.

Equation 8: LP Constraintfor Local Demand

$$\text{Demand}(\text{reg}) \leq Demand_{local}(reg) + Demand_{Import}(reg)$$

Equation 9: LP Constraint for Maximum Local Production

$$\mathrm{Production}_{\max}(\mathrm{reg}_{\mathrm{from}}, pbin) \geq \sum_{reg_{to}}^{Regions} Quantity\left(reg_{from}, reg_{to}, pbin\right)$$

One substantial difference between the LP and NLP approach is the need for constraining the export availability of each region. The NLP approach does not require this constraint since all the coal in a region is represented by a single market clearing price. The use of bins in the LP approach creates a pseudo supply of coal, which is cheaper than market-clearing price. Left unrestrained, the model would use this *cheaper* coal to fulfill demand in other parts of the world to cover for transportation costs. This phenomenon is a direct result of the use of bins to represent a supply curve, combined with a global cost minimization objective function. The constraint to complete the LP problem statement is introduced by explicitly forcing the model to meet local demand before exporting coal resources (Equation 10). Implementing this constraint avoids potentially unrealistic model solutions in which cheaper in-region coal is displaced to meet the demand of other regions at the expense of using more expensive imported coal.

Equation 10: LP Constraint for Satisfaction of Local Demand with Local Production First

$$Export\_Avail(\mathrm{reg}_{\mathrm{from}}, pbin) \geq \sum_{reg_{to}}^{Regions} Quantity\left(reg_{from}, reg_{to}, pbin\right) \forall\, reg_{from} \not\ni reg_{to}$$

Linear Problems have been widely used in EIA forecasting models. A wealth of knowledge has been developed in this area. The LP problem formulation contains a calculated constraint that forces the optimization to satisfy local demand before exports. This constraint generally forces the international market to *see* the marginal price of coal. Otherwise, international coal demand would be satisfied with the cheapest coal to keep global costs down within the problem. The world doesn't work that way.

If a linear fit is not appropriate, a Supply-Price curve can be approximated by using price-quantity bins. Experience with setting up bins demonstrates that the result can be sensitive to bin-construction procedures. This introduces analyst-driven heuristic procedures into the problem.

- Bins too small lead to usage of below-market coal prices.
- Bins too big lead to price-resolution errors and drastic knife-edge effects.

- Heterogeneity in the size of coal supply/demand for countries leads to increasing loss of resolution at the tail ends.

6.4 Model Runtime Execution

In the implementation of the International Coal Market Module, computational operations are divided and organized in a manner that adapts to the looping cycles within WEPS+. These computations in large part represent operations within the Inference Engine and Optimization Engine execution. Computational operations are divided by preprocessor, processor, and postprocessor activities to denote the timing of their execution.

6.4.1 Preprocessor – Data Handling

This category of routines encompass operations to be performed independent of a typical WEPS+ run. These are the sort of computations done in preparation of inputs to WEPS+ associated with the ICMM. Most of the computations within this section are related to the processing and analysis of raw data inputs, which will be stored in a format that can readily be processed by a normal model run. Envisioned computations include:

- Statistical analysis of country information to be used to construct supply-price curves.
- Determination of viable route options for inter-regional transportation of coal. Some transportation options can be readily eliminated from consideration in a model run if the option is found to be too expensive (e.g., transport of coal using Panamax vessels from U.S. East to U.S. West) .
- Other routines necessary for the transformation of third-party data into WEPS+ usable data input.

6.4.2 Preprocessor – WEPS Start

This category of routines encompasses operations that performed at the start of a WEPS+ model run. The routines covered by this category would only run once and would not update as the IEMM gets called by each WEPS+ cycle. Envisioned computations include:

- Reading of external static inputs into model; if possible, model inputs should be kept in-memory to reduce slowdown due to I/O operations.
- User-enabled option for off-loading cycle routines to increase computational speed.

6.4.3 Preprocessor – Cycle Start

This category collects all the routines that would be performed at the start of each cycle of WEPS+. This routine also provides the input interface between the Inference Engine written in Python and the NLP optimization routines in AIMMS. Envisioned computations include:

- Reading of global WEPS+ variables from the HDF common data file.
- Recalculating of Class “A” mine supply points based on macroeconomic feedback and least-squares supply-price curve fit on points. This recalculation could be off-loaded to Preprocessor – WEPS Start at the user’s discretion to minimize computational load.

- Assembly of class B mine supply points based on required macroeconomic independent variables. This recalculation could be off-loaded to Preprocessor – WEPS Start at the user's discretion to minimize computational load.
- Calculation of inter-regional transportation routes matrix based on new fuel pricing information. Routine would select the least-cost option to be used as the transportation cost option.

6.4.4 Processor

This category collects all the routines directly related to the computation of a coal trade solution performed the AIMMS program. This part of the process takes place in the Model part of the Knowledge Based Framework. The aim would be to run AIMMS once per WEPS+ cycle, with yearly iterations to be done as loops within AIMMS program. The problem within AIMMS would be solved on a serial, rolling yearly, sequential basis, taking into account resource depletion, transport capacity expansion, and technology improvements. By warehousing all optimization calculations within AIMMS, the following benefits are achieved:

- Minimization of data exchange needs between Python and AIMMS.
- Solution basis for previous year can be used in the subsequent year to reduce solution search space.
- Rolling basis would enable the use of cumulative constraints, such as reserve depletion and port capacity expansion, and technology advancement implementation.

6.4.5 Postprocessor – Cycle End

This category collects routines used to analyze optimization problem solutions from AIMMS and upload result variables into HDF common data file. This routine also provides the output interface between the Inference Engine written in Python and the NLP optimization routines in AIMMS. This category would also house inter-cycle diagnostic routines to detect solution convergence/oscillation problems and report on individual cycle solve times.

6.4.6 Postprocessor – WEPS End

This final category contains routines that would run after all the cycles have been completed in WEPS+. The purpose of this routine is to finalize all report and debug outputs for the ICMM. The reports to be generated would be to inform the WEPS+ user of the final solution found within the coal trade model and debug reports, which would enable the user to determine the confidence in the solution from the ICMM perspective.

6.5 Rationale for NLP over LP Model Representation

As previously discussed, the use of an NLP framework to solve for the coal trade between regions has been found to be more desirable than an LP approach. The strength of the NLP formulation lies in the use of continuous functions to represent the regional supply-price relationships, which eliminates the emergence of knife-edge problems in model solutions. The NLP formulation also greatly reduces the number of variables the model has to solve.

In an LP formulation, the supply-price relationship would necessarily need to be represented by price-quantity bins. This type of representation introduces discontinuous inflection points, which can lead to model instability. Furthermore, the use of price bins requires careful set-up of prices and quantities, which can become an art unto itself. The Table 5 presents some results obtained by running the sample GAMS code in the appendices. Even though, both model formulations seek to solve the same problem, different solutions were obtained.

Table 5 NLP vs LP Solution Comparison

	Non-Linear Problem	Linear Problem
Quantity of Coal Produced		
Region 1	158.852	180.0 (+13%)
Region 2	108.195	120.0 (+11%)
Region 3	52.953	20.0 (-62%)
Region Market Price		
Region 1	4.426	4.92 (+11%)
Region 2	5.153	5.47 (+6%)
Region 3	6.516	5.530 (-15%)

More work could be done to improve bin size and resolution, which would minimize divergence between NLP and LP solutions; but this by itself increases the modelers' work which would counteract any speed advantage that the LP approach might have.

7 Reporting Routines

7.1 General Approach

Reporting requirements for the ICMM should generally fall into one of three categories: model debug outputs, solution reports, and global variables. The following discussion on the reporting categories can be used by the reporting subjects in the next subsections.

Debug outputs are messages generated by the model that can provide the user an understanding of the model behavior and aid in the identification of potential model errors. The debug model outputs are envisioned to be written in simple text files via the usage of Python's logging functionality. The use of text files as the standard output for debug messages stems from their ease of generation and low resource burden from a programmatic perspective. It is important to note that debug text files can become cluttered if there is no certain hierarchy format. The current system of classification within the logging functionality would be used to aid the user in the identification of problems. The current categories available for logging messages are debug, info, warning, error, and critical[6].

Solution reports cover a category of model outputs that can be considered as formal model outputs, but they are not part of the global variables in the common data file. These solution reports would enable the user to further understand the operation of coal markets within the ICMM. The reports would be generated by creating a PDF file that summarizes and presents the run information. Most of the information contained in the PDF would be generated using openly-available data visualization libraries.

Global variables are those that are shared with other WEPS+ models and are used for the creation of the final WEPS+ report. The global variables for the WEPS+ model are stored within a common file in a Hierarchical Data Format (HDF5). Variables within the common file are updated by other components of WEPS+ until convergence on a global solution is achieved.

The reporting routines described here are for illustrative purposes and should not be considered final. As development of the model is implemented, some reporting routines would necessarily need to be modified for adaptability for different scenarios.

[6] See https://docs.python.org/3.6/howto/logging.html for more information. (Accessed March 2017)

7.2 Supply Curves

The foundation of a solution within the ICMM is the regional supply curves. These curves seek to replicate the production behavior of producers in the face of specific coal prices. As such, it is important for the model to output the final supply curve configuration by region.

These curves are to be treated under the category of solution report. Their mathematical characteristics would be presented as text. The user would be able, via the use of the ICMM control file, determine which model years merit the creation of graphical outputs.

7.3 Transportation Matrix

The final transportation matrix used by the ICMM would be dependent on WEPS+ endogenous inputs from other models. For example, ocean freight costs are dependent on the prices of distillate and residual oil provided by WEPS+. The final transportation cost matrix would represent the lowest transportation cost for a specific route.

The final transportation cost matrix would be contained within the category of solution report. A table will be created containing all the possible region-to-region combinations with their corresponding yearly transportation cost. Additionally, a table would be made available showing the tonnage of coal transported from region to region by year and coal rank. Future enhancements could include further data visualization of the transportation matrix in a GIS layer.

7.4 Optimization Result

Results from the NLP optimization would need to be part of the debug outputs by the model. It is important to check if the optimization routines were able to reach reasonable convergence in a reasonable amount of time. Warning flags would be used to identify conditions indicating calculation time outs. Infeasible model outputs would trigger error flags in the model debug output.

7.5 WEPS+ Level Outputs

Variable outputs to be used by other models would be uploaded to the HDF5 common file. The output of such variables would occur once the model has completed its WEPS cycle computations.

8 Miscellaneous Model Operation Considerations

8.1 AIMMS Run Control

As of the date of this report, no specific Python command and control library has been developed by AIMMS. AIMMS has developed a Software Development Kit (SDK) for several languages (Java, C++, C#, FORTRAN). The SDK allows communication between a program and AIMMS by using computer memory instead of I/O devices. This can speed up operations through bypassing write/read cycles. The SDK also allows for error handling and checking; for example, a license error could initiate timed tries vs. a runtime exception, which might require model shutdown. The SDK could still be used by the development of a *wrapper* in a language like C++ that would communicate with Python.

As a default AIMMS could be run blindly by the Python code using an external process call along with appropriate file input generation. This method has many drawbacks including the inability to distinguish model execution error types and slowdown due to variable transfer operations over the exchange of files.

8.2 Error Handling

A proper error-handling technique permits the user to quickly identify problems with model execution while minimizing the need to trace bugs. A model like WEPS+ is expected to have external users, so it would be in the best interest of users to provide a comprehensive framework for the reporting of model execution errors. Past experience with the National Energy Modeling System (NEMS) indicates that the lack of error handling routines contributes to a reduced expert user base and increases modeling overhead costs. EVA strongly recommends the following error handling category framework as a starting point:

- **Missing or Corrupted Inputs** – This error can be handled by either stopping program execution or continuing with default values. An option could be given to the user to identify location of missing input without killing program execution.

- **License Missing** – Many institutional users of AIMMS use a server-based licensing scheme. Under this scheme, licenses could be temporarily unavailable if other programs or threads are using it. This type of error should lead to a time-delayed re-try of license invocation before stopping the program. This issue becomes sensitive for locations with a need for multiple sensitivity runs.

- **Out-of-Bounds Variables** – Variables deemed as critical would be flagged and tested during model execution. Out-of-bound variables can be an indication of an optimization problem that did not reach convergence or potential user input error. A structured system would be needed where value bounds are classified as trace, debug, notice, warn, error, or fatal. This structure would also permit the shortening of debug outputs.

- **Code Integration Errors** – This type of error should always cause the program execution to stop. This error should contain as much information as possible because a common user would generally need assistance to resolve such a bug.

8.3 Integration with Other WEPS+ Modules

EIA has been in the process of developing new WEPS+ modules in addition to the International Coal Market Module described in this report. A new International Electricity Market Model (IEMM) will provide signals to other models about the consumption of fuels in the electricity sector. The electricity sector is the main consumer of coal in the world, so interactions between the ICMM and IEMM would require special attention.

To improve communications between models, the ICMM should be engineered to provide two sets of outputs. The first set of outputs would be the standard WEPS+ common file parameters that contain coal prices by WEPS+ regions. The translation between coal prices in an ICMM region and a WEPS+ region would be done by demand-weighted aggregation. For example, it is proposed that the United States to be modeled as two regions within the ICMM; but WEPS+ considers the United States to be one region. For reporting purposes, the ICMM would calculate a weighted average by multiplying the price of coal in each region by the demand in each region, sum the output, and divide by the total coal demand in the region, as shown in Equation 11.

Equation 11: Equation for Up-Conversion of Variables from ICMM to WEPS+ Regions

$$Price_{WEPS\ Region} = \frac{\sum\left(Price_{ICMM\ Region} * Demand_{ICMM\ Region}\right)}{Demand_{WEPS\ Region}}$$

The second set of outputs would be a simple comma-separated values (CSV) file that would contain the regional prices of coal as determined by the ICMM by year, region, and rank. This file can be used by other WEPS+ modules, like the IEMM, which would need more detailed information about coal prices. Additionally, the CSV file would contain coal component information such as carbon, sulfur, mercury, and ash content that can be used by pollution control algorithms in other models to determine the generation of pollutants.

Connections between the ICMM and IEMM components of the WEPS+ model still need to be further developed. Discussion about connection of the models have centered on the need for the ICMM to provide price elasticities to the IEMM to reduce model computation times. EVA believes that the use of continuous pricing functions as part of the main coal trade problem statement should significantly reduce the need for model cycling without the need to resort to the export of elasticities.

Nonetheless, if elasticities are needed, EVA proposes such elasticities to be generated by partial differentiation of the resulting objective function with respect to quantities by region. The differentiation of the objective function equation by region should provide a linear slope that can be used by the IEMM to represent coal price elasticities. EVA is available to perform further work on price elasticities should the need arise.

9 Conclusion

EVA's recommends that the International Coal Market Model (ICMM) determine future coal prices and trade flows through the use of quality-adjusted supply-price functions with transportation cost adders. Coal must be treated, at a minimum, as four separate commodities categorized by coal rank (i.e., metallurgical, bituminous, sub-bituminous, and lignite). The quantity and prices for coal by region would be set by the equilibrium of supply functions subject to resource, technological advancement, and infrastructure constraints.

EVA has further found that a Non-Linear Problem (NLP) formulation provides a better representation of coal trade flows than that of a Linear Problem (LP) formulation. Use of an NLP formulation would require the use of convex functions to assure the convergence of the model into a global solution in a reasonable amount of time. The use of an NLP also eliminates the need for mitigation measures presented by LP knife-edge effect solutions.

The final ICMM should be developed by using a Knowledge Based Framework (KBF); this would entail the addition of raw data conversion routines into the ICMM. All routines for the ICMM would be grouped by their respective start/stop times during a model run.

This CDR also provides a set of recommendations for AIMMS run control from Python and Error Handling. Both sets of recommendations would enable a WEPS+ user to better understand any problems that arise during model execution.

10 References

Boyd, S., & Vandenberghe, L. (2004). *Convex Optimization.* Cambridge University Press.

Hellerworx. (2016). *DOE/EIA International Coal Supply Curve and Trasportation Modeling Australia Pilot Study, Ocean Freight Forecast & Coal Data Survey Study Methodology and Documentation.* Washington D.C.: Energy Information Administration.

International Energy Agency. (2016). *Coal Information.* Paris: International Energy Agency.

Petromindo.com. (2016). *Indonesia Coal Book.* Jakarta: Petromindo.com.

11 Appendix: GAMS NLP Experimental Model Formulation

```
$Title Example of Proposed NLP Construction for Int'l Coal CDR

*NLP Problem Parameter and Set Setup
Set reg Local Coal Regions /REG1,REG2,REG3/;
Alias(reg,from_reg);
Alias(reg,to_reg);

Parameter Price(reg) Base Coal Prices by Region
    /REG1 2.0
    REG2 3.0
    REG3 5.0/;

Parameter Price_Elas(reg) Price Elasticity for Production
    /REG1 0.005
    REG2 0.005
    REG3 0.005/;

Parameter Max_Q(reg) Maximum amount of Coal that can be produced in Region
    /REG1 1000
    REG2 1000
    REG3 1000/;

Parameter L_Demand(reg) Local Coal Demand
    /REG1 20.0
    REG2 100.0
    REG3 200.0/;

Table T(from_reg,to_reg) Coal Export Transportation Cost Table
            REG1   REG2   REG3
     REG1   0.0    0.3    0.3
     REG2   0.3    0.0    0.3
     REG3   0.3    0.3    0.0;

*Varible and Equation Declaration
Variable t_cost total system cost;

Positive Variables
    P_Supply(reg) Coal Mining Cost
    Q(from_reg,to_reg) Production Quantities
    Q_Local(reg) Production for local use
    Q_Export(reg) Production for export
    Q_Import(reg) Production for import;

Equations
    supply_price Coal Supply Price Function
    cost Objective Function with Total System Cost to be Minimized

    imports(reg) Imports to Region
    exports(reg) Exports out of Region
    local_prod(reg) Production for Local Use

    demand(reg) Local Coal Demand
    max_prod(reg) Maximum Coal Production by Entity;

*Objective Function:
supply_price(reg).. P_Supply(reg) =e= Price(reg) * exp(sum(to_reg, Q(reg,to_reg)) * Price_Elas(reg));
cost.. t_cost =e= sum((from_reg,to_reg), Q(from_reg,to_reg) * (P_Supply(from_reg) + T(from_reg,to_reg)));

*Flow Equations
local_prod(reg).. Q_Local(reg) =e= sum((to_reg), Q(reg,to_reg)$(ord(reg) = ord(to_reg)));
exports(reg).. Q_Export(reg) =e= sum((to_reg), Q(reg,to_reg)$(ord(reg) <> ord(to_reg)));
imports(reg).. Q_Import(reg) =e= sum((from_reg), Q(from_reg,reg)$(ord(from_reg) <> ord(reg)));
```

```
*Quantity Constraints
demand(reg).. L_Demand(reg) =l= Q_Local(reg) + Q_Import(reg);
max_prod(reg)..   Max_Q(reg) =g= sum((to_reg), Q(reg,to_reg));

*Problem Solver Settings
Model intlcoal /all/;
Option nlp = conopt;
Solve intlcoal using nlp minimizing t_cost;

*Result Reporting Routines
Parameter q_result(reg);
q_result(from_reg) = sum(to_reg, Q.l(from_reg,to_reg));

Alias(reg,i);
Parameter p_result(reg);
Loop(i,
      p_result(i) = Price(i) * exp(sum(to_reg, Q.l(i,to_reg)) * Price_Elas(i));
);

Display q_result;
Display p_result;
```

12 Appendix: GAMS LP Experimental Model Formulation

```
$Title Example of Proposed LP Construction for Int'l Coal CDR

*LP Problem Parameter and Set Setup
Sets
    reg Local Coal Regions /REG1,REG2,REG3/
    pBins Coal Supply Curve Price Bins /r1*r10/;

Alias(reg,from_reg);
Alias(reg,to_reg);

Table Price(reg,pBins) Coal Price by Bin
          r1    r2    r3    r4    r5    r6    r7    r8    r9    r10
    REG1    2.32  2.70  3.14  3.64  4.23  4.92  5.72  6.64  7.71  8.96
    REG2  3.32  3.66  4.05  4.48  4.95  5.47  6.04  6.68  7.38  8.15
    REG3  5.53  6.11  6.75  7.46  8.24  9.11  10.07  11.13  12.30  13.59;

Table Max_Q(reg,pBins) Coal Max Quantity by Bin
          r1    r2    r3    r4    r5    r6    r7    r8    r9    r10
    REG1    30.0  30.0  30.0  30.0  30.0  30.0  30.0  30.0  30.0  1000.0
    REG2  20.0  20.0  20.0  20.0  20.0  20.0  20.0  20.0  20.0  1000.0
    REG3  20.0  20.0  20.0  20.0  20.0  20.0  20.0  20.0  20.0  1000.0;

Parameter L_Demand(reg) Local Coal Demand
    /REG1 20.0
    REG2 100.0
    REG3 200.0/;

Table T(from_reg,to_reg) Coal Export Transportation Cost Table
          REG1   REG2  REG3
    REG1    0.0    0.3    0.3
    REG2  0.3    0.0    0.3
    REG3  0.3    0.3    0.0;

*Determination of Max Coal Exports, Assumes Local Demand is Met First
Parameter Export_Available(reg,pBins) Amount of Export Available by Bin After Local Demand;
Alias(reg,i);
Alias(pBins,j);
LOOP(i,
    LOOP(j,
        Export_Available(i,j) = sum((pBins), Max_Q(i,pBins)$(ord(pBins) <= ord(j))) - L_Demand(i);
        IF(Export_Available(i,j) < 0.0, Export_Available(i,j) = 0.0);
        IF(Export_Available(i,j) > Max_Q(i,j), Export_Available(i,j) = Max_Q(i,j));
    );
);
Display Export_Available;

*Varible and Equation Declaration
Variable t_cost Total System Cost;

Positive Variables
    Q(from_reg,to_reg,pBins) Production Quantities
    Q_Local(reg) Production for local use
    Q_Export(reg) Production for export
    Q_Import(reg) Production for import;

Equations
    cost Total System Cost to be Minimized

    imports(reg) Imports to Region
    exports(reg) Exports out of Region
```

```
    local_prod(reg) Local Production

    max_prod(reg,pBins) Maximum Production by Price Bin
    min_local_prod(reg,pBins) Local production over exports constraint
    demand(reg) Local Coal Demand;

*Objective Function:
cost.. t_cost =e= sum((from_reg,to_reg,pBins), Q(from_reg,to_reg,pBins) * (Price(from_reg,pBins) + T(from_reg,to_reg)));

*Flow Equations
local_prod(reg).. Q_Local(reg) =e= sum((to_reg,pBins), Q(reg,to_reg,pBins)$(ord(reg) = ord(to_reg)));
exports(reg).. Q_Export(reg) =e= sum((to_reg,pBins), Q(reg,to_reg,pBins)$(ord(reg) <> ord(to_reg)));
imports(reg).. Q_Import(reg) =e= sum((from_reg,pBins), Q(from_reg,reg,pBins)$(ord(from_reg) <> ord(reg)));

*Quantity Constraints
*Demand <= Local Production + Imports
demand(reg).. L_Demand(reg) =l= Q_Local(reg) + Q_Import(reg);
max_prod(reg,pBins)..   Max_Q(reg,pBins) =g= sum((to_reg), Q(reg,to_reg,pBins));
*Local Export Available >= Local Production Exports
min_local_prod(from_reg,pBins).. Export_Available(from_reg,pBins) =g= sum(to_reg, Q(from_reg,to_reg,pBins)$(ord(from_reg) <>
ord(to_reg)));

*Problem Solver Settings
Model intlcoal /all/;
Solve intlcoal using lp minimizing t_cost;

*Result Reporting Routines
Parameter q_result(reg);
q_result(from_reg) = sum((to_reg,pBins), Q.l(from_reg,to_reg,pBins));

Parameter p_result(reg);
Loop(i,
    Loop(j,
        If(sum(to_reg, Q.l(i,to_reg,j)) > 0, p_result(i) = Price(i,j);)
    ;)
);

Display q_result;
Display p_result;
```