

**GLOBAL HYDROCARBON
SUPPLY MODEL
Hydrocarbon Gas Liquids
Component Design Report**

Prepared for

**Office of Integrated Analysis and Forecasting
Energy Information Administration
U.S. Department of Energy**

Prepared by

Anne B. Keller
Midstream Energy Group, Inc.
10707 Corporate Drive
Suite 158
Stafford, TX 77477
(281) 277-5560

Table of Contents

1. Overview
 - a. Purpose of the Study
 - b. Rationale for Improvements to Existing Models - GHySMo
 - c. Purpose of the Hydrocarbon Gas Liquids (HGL) submodule
2. Conclusions and Recommendations
 - a. Energy Balances – Modelling Issues
 - b. Database Design Considerations – Use of Detailed Attribute Definitions
 - c. Use of Nodes for Point Source Supply Aggregation and Yields Models
 - d. Build Systematic Validation and Calibration Processes into the Module Design
3. Methodology description
 - a. Model objective
 - b. HGL Supply Module Design Basis - Overview
 - c. Resource Representation - Extraction Process Supply Channel
 - d. Resource Representation - Conversion Process Supply Channel
 - e. Model Design Recommendations
 - f. Data requirements
 - g. Software considerations
4. Uncertainty and limitations
5. References - Definitions

Overview

Purpose of the study

This document presents a proposal for a new model for enhanced supply forecasting for a growing portion of the fuels market – liquids extracted by processing hydrocarbon streams that are typically in vapor phase at atmospheric temperature and pressure. This model would provide forecasts of global Hydrocarbon Gas Liquids (HGL) production from natural gas field processing, stand-alone condensate splitting, and coal/gas/biomass syngas to liquids conversion processes (xTL). These forecasts can then interact with or become inputs to other models of processing, transport, distribution, and storage of natural gas and liquid fuels.

Rationale for Improvements to Existing Models – GHySMo Project

The Office of Petroleum, Natural Gas, and Biofuels Analysis (PNGBA) of the EIA has been tasked to develop a dynamic representation (referred to here as GHySMo¹) of the global production, processing, transport, distribution, and storage of natural gas and liquid fuels. The primary function of this model will be to replace portions of the existing upstream and midstream models of petroleum and natural gas within the World Energy Projection System Plus (WEPS+)². The outputs of this system are used to produce EIA's International Energy Outlook. The ultimate purpose of this project is to improve EIA's capability to represent international markets (i.e., prices and commodity flows) for liquids and natural gas under a variety of assumptions.

EIA recognizes that the development of GHySMo is an ambitious project. As a multiple-purpose analytic platform, GHySMo must retain flexibility in inputs, operational modes, and outputs. This will require a sophisticated understanding and management of data, analyst judgment, and other forms of knowledge (such as what is and is not known, and levels of certainty of the input data). EIA strongly suggests that the model and the data feeding the model be managed separately as much as possible, so that as data sources change, the model logic and code can remain minimally affected. The model should impose requirements on the knowledge management (KM) system, and the KM system should seek to satisfy those requirements as best as possible given the available data.

WEPS+ has component modules for other areas of the energy system, particularly demand modules, and a macroeconomic module that will provide regional, sector-specific annual consumption estimates for natural gas and liquids products, and that will be provided corresponding wholesale and/or delivered product prices by GHySMo. GHySMo will need to also produce supply prices and quantities and trade volumes and prices for publication in the IEO.

Another secondary function of GHySMo is to operate in a standalone fashion. Without the additional overhead of an integrating framework, it is hoped that a standalone version of the model will run dramatically faster. Not only will this improve testing and development turnaround times, but it should enable the targeted use of greater levels

¹ The Global Hydrocarbon Supply Model, which is generally assumed to be a computational model throughout this document, but which is understood to include any representation of the global liquids and gas markets that meets the requirements set forth in this document.

² <http://www.eia.gov/forecasts/ieo/models.cfm> WEPS+ is a modeling system used to generate the projections published in EIA's International Energy Outlook (IEO).

of detail to support certain topical analyses that may not critically depend on dynamic feedback from outside the liquids and gas markets. It is envisioned that a standalone GHySMo would be used to perform such analyses as deep-dive analyses of specific countries or regions of the world, changes to specific national policies or pricing strategies, the impact of new fuel standards or formulations, what-if scenarios involving specific pieces of infrastructure, and other topical analyses as the need arises.

Purpose of the Hydrocarbon Gas Liquid (HGL) Module

Energy is a key, if not THE key, input which supports economic growth. An increasing amount of the primary energy we consume today, and are planning to consume in the future, comes from the lighter portion of the hydrocarbon stream. This portion of the hydrocarbon energy supply portfolio is comprised primarily of methane through hexane molecules which generally physically present themselves in vapor form at standard atmospheric temperature and pressure. Whereas crude oil is in a liquid state under surface conditions, and requires energy for pumping, and heating for initial separation into intermediate products, the production of Hydrocarbon Gas Liquids (HGL) requires the consumption of considerable energy to provide compression and, depending on end use demand, refrigeration, to separate the streams into their component parts and transport them to end use markets.

As production of high gravity oils continues to grow, there have been a number of discussions focused around the issue of modelling condensate supplies. In the past, these streams, which are usually described as hydrocarbons having an API gravity of 50° or higher, have typically blended into crude oil and transported in the logistics systems that handle crude. The production volumes are now large enough that blending them produces a crude oil that is not suitable for processing in many refineries due to the relatively high naphtha content and relatively low resid and distillate content compared to the crude oil that most plants are designed to process. In addition, the refined product yields from these streams are not as profitable as those from more balanced crudes. The physical state of condensate, which is liquid at atmospheric temperature and pressure, places it in the realm of oils rather than HGLs. However, some volumes of a liquid that has the same characteristics as “pentanes plus” extracted from gas in a gas processing facility often ‘fall out’ of the gas vapor stream in the gathering system between the wellhead and the gas plant due to temperature (cold weather), or pressure (going through compressors to boost the pressure in a gathering line). These streams are NOT considered to be HGL’s in that they are not the result of on purpose extraction processes. In other words, the “pentanes plus” volumes for which this module is producing supply projections are those which remain in the gas stream all the way into the inlet of the gas processing facility and are extracted in the process train.

Condensate streams are usually described based on where they are first produced, because in the United States the revenue from the sale of these products is often based on who has paid for the asset that captures them:

- 1) Lease Condensate – volumes which emerge on the surface from the wellhead in liquid form, but do not have sufficient heavier molecules to achieve a gravity rating to be classified as crude oil. These barrels are usually captured near the wellhead in lease separators which are designed to remove water from the liquid hydrocarbons. The so-called “liquefiable” hydrocarbons which remain in the gas stream are captured as vapor and moved through the gas gathering system.
- 2) Field Condensate – volumes which emerge as liquid between the lease or so-called Central Delivery Point where a number of wellhead flowlines come together for measurement and testing and the inlet of a gas plant or the sales meter for the gas pipeline. These volumes are usually produced during compression. They are collected in a field tank and may be trucked to a crude oil terminal for blending and sale as crude or to a tank located at a gas plant for injection into an NGL pipeline or trucking to a

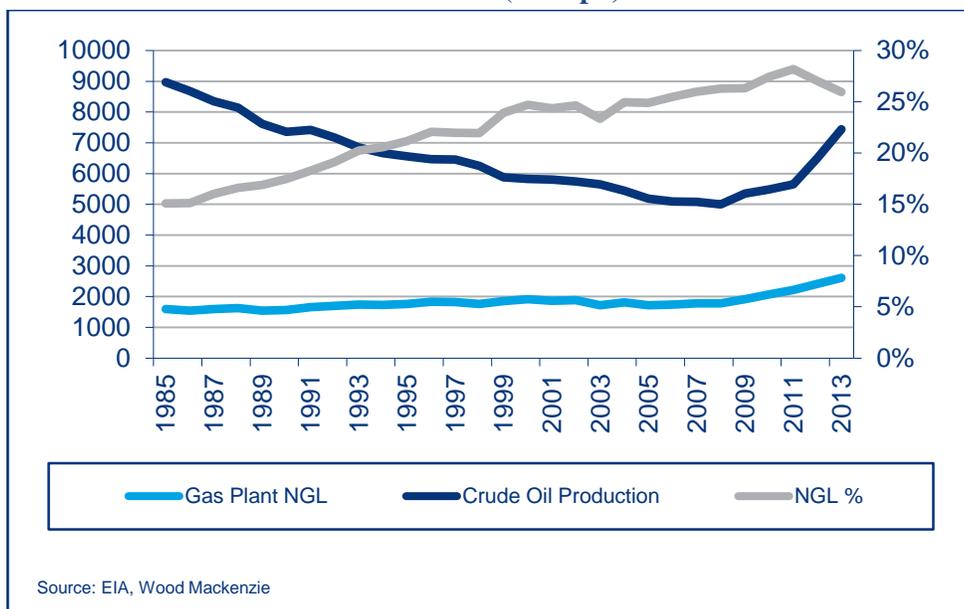
collection point for sales. If the same company operates the gathering system and the gas plant or has a common ownership in them it is often more cost effective to co-mingle the field condensate with the plant condensate (see next definition) for measurement and transportation to market.

- 3) Plant, or Inlet condensate – volumes which, like field condensate, emerge in front of the processing train at the gas processing facility as a liquid. These volumes are usually produced as a result of compression which occurs near the inlet of the gas processing facility. Gas plant operators often boost the pressure of the gas coming into their plants in order to maximize the effectiveness of a mechanical process which increases HGL recovery without the use of refrigeration. This process, called turbo-expansion, passes high pressure gas through an expander (basically wide place in the pipeline), reducing the pressure by several hundred PSI instantly, creating a cooling effect which drops out NGL's. These volumes often contain enough heavier HGL components to be blended into crude and sold separately.

None of these condensates would be considered an HGL as defined in this module, since they were produced prior to the introduction of the gas stream into a processing train specifically designed to liquefy and separate hydrocarbons from methane. This admittedly fine distinction likely developed due to contract terms which allowed the owner of the compressor stations to 'save and sell' the liquids resulting from compression for the owner's account, whereas the revenues from liquids removed from a gas stream in a processing plant would usually be shared with the producer.

As an example of the growing importance of HGL's in the energy portfolio, the rapid growth in US oil and gas production from shale formations over the past decade has increased the percentage of HGLs supplied as part of the overall hydrocarbon stream, as shown in Figure 1.

**Figure 1. US Crude and Gas Plant NGL Production
1985-2013 (000 bpd)**



The available volume and the current difference in relative values between oil and gas based fuels appears to support the significant investments required to consume more of this gas-based energy in liquid form, in baseload applications such as petrochemical production, as well as supplying additional volumes to global trade

well as for other EIA projects. These are summarized in this section and are incorporated in the recommended approach to structuring the work process and models described in detail in the body of the report.

Energy Balances – External Issues That Impact the HGL SubModule

In order to better ensure that the level of economic activity in the overall energy forecast can be supported as fuel supply sources shift over time, we recommend that a version of the “mass and heat” balancing approach that process plants use to evaluate their operations be adopted as part of the macro-economic module that provides the demand portion of the HGL sub-module assumptions. In order for all the sub-modules to properly respond to the economy’s ‘call’ on energy, the macroeconomic module should incorporate a process by which overall global energy requirements are determined first based on the assessment of economic growth projections, and assign these requirements to sectors.

Inclusion of a mechanism to provide a better measure of price sensitivity by sector would also be useful as a way to allocate energy resources by the appropriate mechanism, whether it is cost of production, cost of substitution (relative value in use), or value of conservation. This process should allow for the substitution of fuels with differing energy densities, and for the alternate value of fuels as feedstock for industrial products. Examples include ethane and propane used for ethylene production instead of heating, natural gas for electricity generation to provide fuel for transportation, and ethanol for normal butane and MTBE in motor fuels.

Database Design Considerations – Use of Detailed Attribute Definitions

Attribute definitions for data elements included in configuring the model must be described in sufficient detail for all users to understand and interpret the same way to allow calibration across multiple markets and common understanding of the context in which conclusions are reached. This is usually called ‘normalization’ in systems design documents.

Data elements should be captured at a level of detail that will allow the model to be fit for purpose over a long period of time without requiring revisions when political boundaries shift. To that end, some modifications or additions to the NEMS data tables are recommended to accommodate a lower level of detail within some product definitions. Alternatively, the detailed outputs from the HGL module can be summed to a level that is usable within the current NEMS data structure.

Use of Nodes for Volume Aggregation

Advanced technology and programming can support a shift to a nodal model that focuses first on estimating supply from defined sub-surface resources, then aggregates the volumes into sub-regional and regional "supply" and "market" zones for input into logistics models for trade flow balancing and further analysis. Using nodes also normalizes the raw input data into the building blocks that are needed for other calculations, such as trade balances and logistics costs calculations.

Build Systematic Validation and Calibration Processes Into the Module Design

The workflow and model design structure should incorporate methods for comparing model results to external data and/or observed market conditions to serve as a “teaching tool” to improve the model’s accuracy over time. A triangulation process whereby the output of the HGL model is compared to similar outputs from other sources, such as trade statistics, engineering papers, industry surveys, and National Oil Company (NOC) data sources, and surveys conducted by government agencies such as the EIA itself, will be useful in identifying input assumptions which require review and revalidation before the next model run. In regions where data sources are incomplete or found to be unreliable, market events and price movements can serve as secondary indicators in assessing the reliability of the modelled results.

Methodology Description

This section of the report describes the model's objective, proposed structure, and other attributes related to the design basis for the model. This description is at the level of a functional design document, which serves as a guide for the development of the detailed design plan to produce the module.

Model Objective

The objective of the proposed model is to provide a means of estimating and forecasting the supply of Hydrocarbon Gas Liquids (HGL) produced outside the US. The NEMS model currently produces projections of US HGL supply that can continue to be used for US regions. If a decision is made to use a uniform methodology, this module can generate the supply projections for NEMS as well. This new model is to enhance supply forecasting for a growing portion of the fuels market – liquids extracted by processing hydrocarbon streams that are typically in gaseous phase at atmospheric temperature and pressure. This model would provide forecasts of global Hydrocarbon Gas Liquids (HGL) production from natural gas field processing, condensate splitting, and coal/gas/biomass syngas to liquids conversion processes (xTL). These forecasts can then interact with or become inputs to other models of processing, transport, distribution, and storage of natural gas and liquid fuels.

The bright line test to distinguish between the supply forecasts generated by this module and the supply forecasts generated by refinery modules such as PMM is that the assets included in the HGL model use process technologies which do not involve intermediate conversion steps or blending. Crude oil refineries would be considered demand sources for HGLs based on their price forecasts ("value in use" assessments) for HGL components. Their projected production of HGL's (propane and butane, for example) would be developed using refinery models and these volumes would be inputs to the overall supply HGL supply balances for this module.

HGL Supply Module Design Basis – Overview

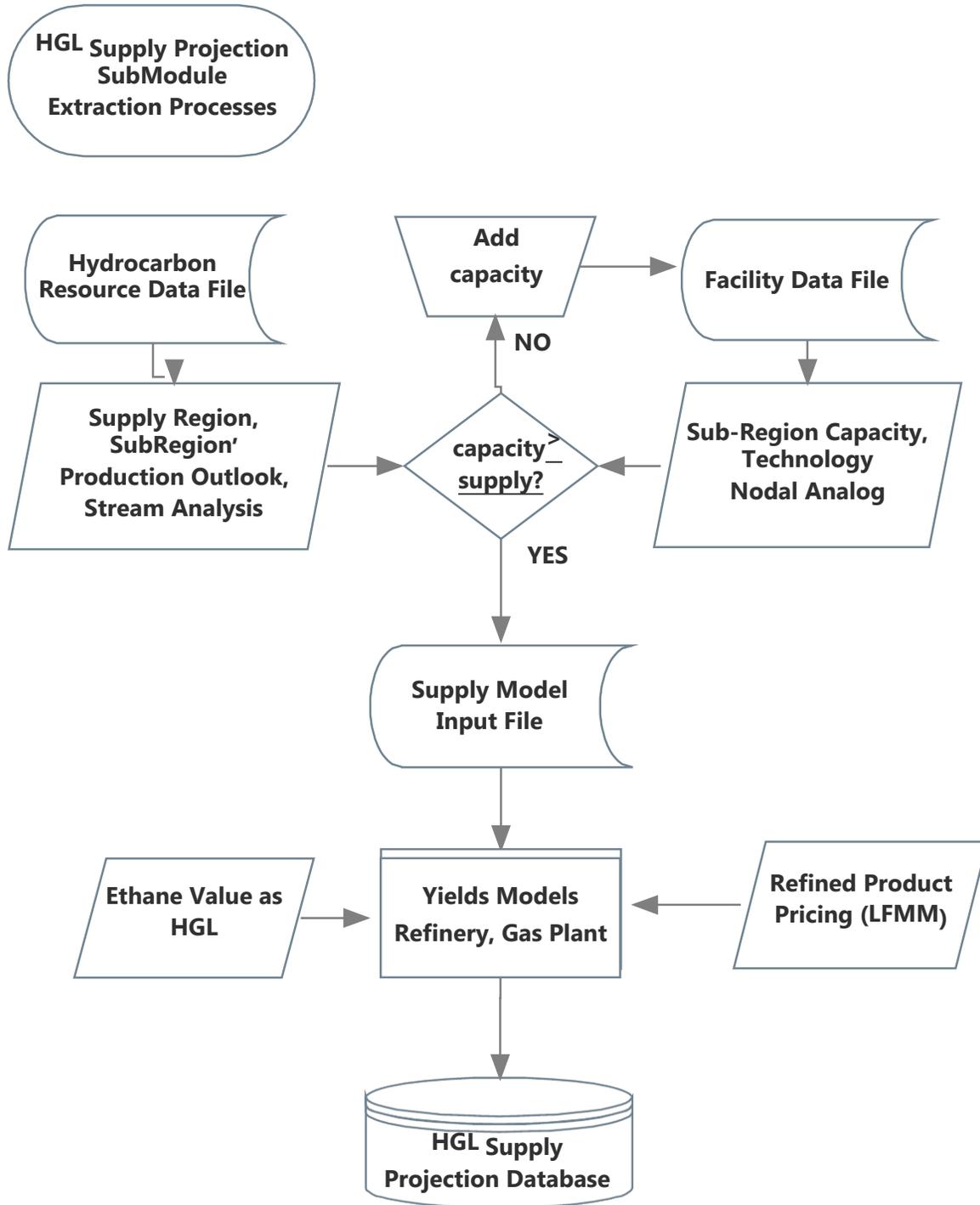
The proposed design consists of a series of components - databases for model inputs and results, and programs which perform analysis and optimization functions. These components can be built and maintained independently, which allows data maintenance and design changes to be made to portions of the module without disturbing the overall "modelling ecosystem" of the module. This design basis is also more cost effective in that less complex tasks, such as updating location and asset capacity and technology information and posting of historical results for benchmarking, can be more easily defined, outsourced, and managed.

The primary unit of organization for the HGL supply module is based on the mode of production, which sub-divides the HGL asset groups comprising the supply sources into 2 channels. One is based on extraction, one on conversion. The extraction supply channel consists of the assets that produce HGLs via extraction – gas processing facilities using absorption, compression, and/or refrigeration to condense HGLs from an incoming vapor stream, including those which remove NGLs from natural gas prior to liquefaction for export as LNG. These plants are also located in refineries – the HGLs they produce are categorized as sourced from oil refining as opposed to gas processing and are entered as inputs from the refinery models. The conversion supply channel comprises assets that convert methane to liquid fuels – gas to liquids, coal to liquids, methanol to liquids, and biofuels.

Resource Representation – Extraction Process Supply Channel – Subsurface Node Definition

The first step in developing HGL supply projections is to define the universe of hydrocarbon resources that will be included in the analysis (sometimes described as sub-surface "assets" by industry analysts). A flow diagram of the data and processes recommended to be included in the supply projection portion of the module is provided at Figure 3.

Figure 3. HGL Supply Projection – Volumes From Extraction Processes – Supply Modelling Workflow



The recommended regional and sub-regional levels of volume aggregation are described in the following section.

Level 1 Summary – Minimum of 7 Regions

Since geopolitical boundaries within continents could shift over the projection time horizon, the proposed 'asset' database or catalog of hydrocarbon reserves is proposed to be summed into the same overall regions currently used by the EIA in its current global analyses¹, with some shifts to group countries into trade relationships as opposed to political affiliation. The 7 regions are:

- 1) North America – United States, Canada, and Mexico
- 2) Middle East and North Africa – includes Turkey due to trade flows
- 3) Asia and Oceania – includes Australia and India
- 4) Western Europe
- 5) Central and South America
- 6) Sub-Saharan Africa
- 7) Eurasia – includes the Former Soviet Union and Eastern Europe

If conformity with the existing International Natural Gas Model (INGM) is desired, the regional aggregations can be expanded to include the 16 regions currently in this model. This would allow the gas supply volume projections from this model to be used as starting points for the determination of HGLs without requiring further summation.

Level 2 – Regional Hydrocarbon Sources

Within these 7 (or 16 if using the INGM regional aggregation scheme) regions, resource definition should be captured in the dataset to define the major oil and gas reserves within each of the regions. At the individual 'play' and 'sub-'play' level, which are typically defined in terms of specific portions of reservoirs, the global universe is immense – several thousand items. This level of detail is unnecessary for HGL projections. The cost and expense of removing the HGL streams from gas limits the number of areas that need to be included in the HGL projection module to those which are capable of producing enough gas to justify the cost of transporting it. Further, HGL from oil is typically removed in refineries, so the primary input for determining levels of HGL extracted as part of the oil refining process can be captured at the facility level via estimates of throughput into refineries and splitters as opposed to tracking oil production volumes at the field level.

The final list of sub-regions will depend on the threshold of materiality determined by the modeling team to be acceptable. For efficiency, these should ideally be the same resources that are captured in the modules which provide the inputs to the primary hydrocarbon supply projections (oil, natural gas) for the overall energy balances, currently the International Natural Gas Model. The 2012 world map of energy reserves on the EIA website at <http://www.eia.gov/countries/index.cfm> shows that most of the world's proved oil reserves are located in only 17 countries. According to the EIA's international energy statistics for dry gas production, (see <http://www.eia.gov/cfapps/ipdbproject>), there are only 33 countries which reported natural gas production above 500 Mmcf/day. An approach that focuses on projections for the sub-regions comprising 80% of total worldwide gas production would need to track facility details only for North America, Russia, the MidEast, China, and the LNG liquefaction plants located outside these countries.

Level 3 - Sub-Regional Resource Allocations

The use of additional sub-regions (Level 3) to divide Regional Resources among geopolitical zones (Level 2) provides the ability to shift geopolitical boundaries within a region - for example, to allocate production in areas with shared resources, such as the South Pars (Iran) and North (Qatar) Fields, which overlap the offshore boundary between these nations. The respective national oil companies within each country control the supply of HGLs within their borders, and the facilities used to extract them are located onshore within each country's current boundaries. Although both geopolitical units currently fall within the Middle East North Africa (MidEast_NA) resource region 2, the use of a volume allocation mechanism provides a means of compiling supply projections that can be compared to reported volumes for each country to validate and calibrate the projections in the historical models. As an example of the importance of this approach, Qatar and Iran are in the top 10 list of countries with hydrocarbon resources, so the volumes involved can be material to each country's energy balance.

This approach can also be used for North America to divide supply across regions. Examples of major North American resources which require allocation include the Burgos Basin (US PADD III/Mexico), where the US Eagle Ford play overlaps the Mexican border, the Montney Shale and Horn River Basin (British Columbia and Alberta Provinces), and the Bakken Shale (United States and Canada). Canada can also be divided into Eastern and Western Regions for purposes of modelling cross-border transfers and waterborne trade.

Table 1. is an example of the levels of aggregation to be captured in developing the Regional Resources database.

Table 1. Hydrocarbon Resource Database
Levels of Aggregation for Supply Projection Information - Example

Region	Level 1 Region Name	Level 2 Resource Name (Basin/Field)	Resource Type	Level 3 Sub Region Name1	Level 3 Sub Region1 Controlled %	Level 3 Sub Region Name2	Level 3 Sub Region2 Controlled %
2	MidEast_NA	South Pars Field	Oil	Iran	Xxx	Qatar	Xxx
2	MidEast_NA	Ghawar Field	Oil	Saudi Arabia	100		
2	MidEast_NA	Safaniya Field	Oil	Saudi Arabia	100		

Resource Representation – Extraction Process Supply Channel – Facilities Database

The next step in building the HGL supply module is to develop the initial dataset of surface facilities which are (or will be) used to extract the HGL's from the produced oil and natural gas streams. These are generally described as gas processing plants, refineries, 'conditioning' plants, and condensate splitters.

Facilities that handle gas are generally located closer to the producing area than are refineries, which tend to be located closer to the market. In case of gas, it's often necessary to remove sufficient HGL from the gas stream to allow the gas lines to operate efficiently and safely. For oil, it's usually less costly to transport the crude to a location closer to where the products will be used or distributed to other markets since there is only one stream to transport rather than numerous refined products.

This dataset will require at least one common descriptor field to allow the Hydrocarbon Asset database to be joined to the Facilities database to associate infrastructure with projected oil and gas production volumes. Gas processing plants should be associated with fields or basins in order to determine whether capacity is sufficient to handle the projected production throughout the forecast horizon. Condensate splitter capacity can be compared on a sub-region basis since since condensate is generally transported to storage and distribution center rather than being processed in the oil field. The database contains the fields that allow this association to be included as needed to build the models.

Table 2. provides suggested data fields for the Facilities database – the fields from the Hydrocarbon Regional Resource Database can be joined to this database using the Level2 or Level3 definitions to associate them with the proper source of production volumes to be used in generating the supply projections in the Supply Projection Input file.

Table 2. HGL Supply from Extraction - Facilities Data Descriptions
Example Facility Detail Database file segment

Region ID	Level 1 Region Name	Logistics Location Geocode	Level 2 Resource Name	Facility_ID	Facility Name	Facility Type	Capacity Units	Capacity
2	MidEast_NA	Arab Gulf Jubail	Ghawar Field	XXX	Hawiyah	Gas Processing Plant	Mmcf/d	2,400
2	MidEast_NA	Red Sea Yanbu		XXX	Rabigh	Condensate Splitter	Bpcd	550,000

Table 2. Field Descriptor Definitions:

"Facility Type Code" - a code to identify each type of facility (condensate stabilizer/splitter, gas processing plant, LNG plant, etc.). Use of a code table reference system allows the fields to be uniformly populated for queries since the names only need to be modified in the code table instead of the datafile itself.

“Logistics Location Geocode” – is used to link the supply projections to logistics costs datasets for purposes of determining the regional value of products produced. This is usually the closest point where volumes are aggregated for transfer out of the field or sub-region via pipeline or ship. This can be a latitude/longitude point, the name of the closest port, an injection point or group of points on a pipeline system, or a rail or truck terminal location depending on the sub-region and the data available in the logistics files. The group of facilities having the same facility type and Logistics Location Geocode is the logical starting point for supply/facility nodal definitions.

It should be noted that capital costs and operating expenses are not included in the extraction facility database – HGLs other than ethane above a minimum level required to meet pipeline specifications are considered by-products of gas processing and oil refining, and their prices are derived from the buyer’s alternative for the end use contemplated as opposed to the supplier’s cost of production. In other words, if the gas contains enough HGL to require processing, the supplier is a price-taker for the volumes which must be recovered to meet pipeline Btu requirements.

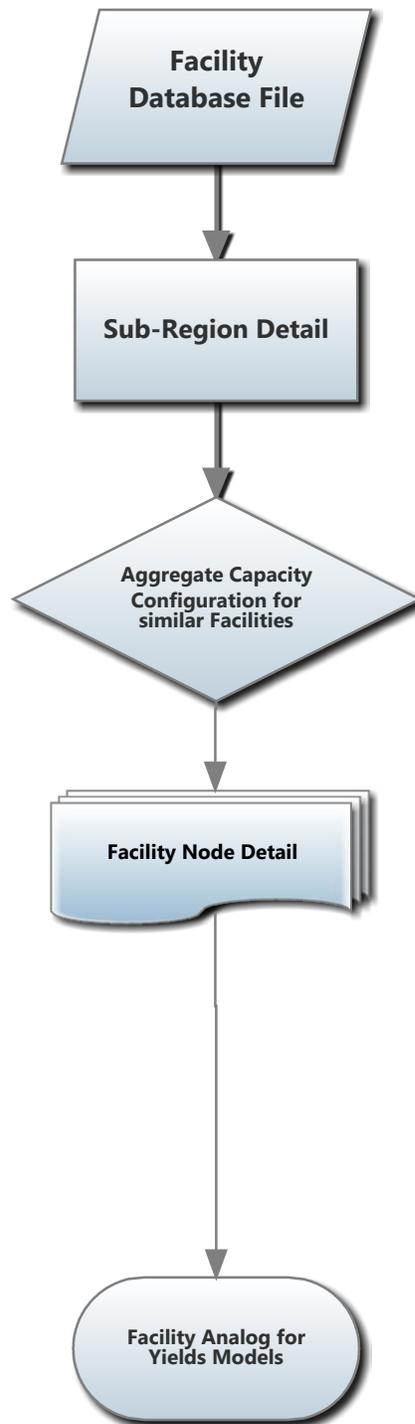
Resource Representation – Extraction Process Supply Channel – Facility Node Descriptions

Instead of modelling supply at the individual facility level, one or more "Facility Analog Models" are selected from a group developed for the yields model as representative of the facilities in the sub-region. These analogs are summaries of the capacity and configurations of the actual plants for which operating data can be obtained to use as a proxy. As noted above, the groupings will be apparent once the facility locations with respect to takeaway logistics have been reviewed. The data to use initial system configurations and test runs can be developed by modeling plants for which reported actual yields are available, by reviewing reported yields in survey data such as the Oil and Gas Journal information, which includes capacity, technology, and estimated production data from plants worldwide, from technical papers published by firms specializing in process technology, and from information supplied by producers in the region.

If there are a number of different process technologies in use which result in significant variations in yields (for example, having cryogenic and absorption gas processing plants in the same region or sub-region), it may be necessary to define more than one node within a region or sub-region. Generally speaking, an acceptable result can be reached without having more than one facility analog standing in for the individual plants in a sub-region. Exceptions will likely be in sub-regions where legacy assets are in place to handle volumes from fields that have been in production for many years alongside new facilities developed to handle supply from newer wells to service growing downstream demand. Given that there are relatively few countries outside the US and Canada with multiple gas processing plants, there should not be many instances where this level of analysis is necessary.

Figure 4. provides an overview of the process of defining a facility node for use in producing supply yields estimates.

Figure 4. HGL Supply Module – Facility Database Nodal Capacity Aggregation
Facility Node Definition



Resource Representation – Extraction Process Supply Channel – Production Volumes (Gas)

After the Hydrocarbon Resource database and the Facilities database have been populated, it will be necessary to estimate gas production volumes to be used in the gas processing yields models, and the crude throughput volumes to be used to calculate refinery HGL yields.

The optimal method for acquiring the initial supply datasets to create the projections would be to obtain a datafile from a service which produces these forecasts for a global client base. IHS and WoodMackenzie have global datafiles and the resources to maintain them, but their services may be beyond the budget available for this module. The EIA has already developed its own projections for US oil and gas production. Projections for Canadian production are readily available on the Canadian Association of Petroleum Producers (CAPP) website, from CERI (the Canadian Energy Research Institute), and from the Canadian National Energy Board. Outside the US and Canada, organizations such as the IEA, Cedigaz, National Oil Company (NOC) websites, and project updates and energy surveys produced by private companies such as Shell, Exxon, ENI, and BP can provide information to use in assessing production. Although it is possible to transport gas and process it in another country, the facilities involved are located at LNG receiving terminals, and information around their HGL production capabilities and throughputs is available via industry resources such as the *LNG Journal*, *Gas and Power*, and other trade publications.

Resource Representation – Extraction Process Supply Channel – Plant Inlet Volumes

Due to the expense of constructing, operating, and maintaining gas processing plants, some gas may not be processed to extract HGLs. The projected production volumes will need to be adjusted to produce an estimated of the gas volumes that will be moved through processing facilities – also known as ‘inlet volumes’. This is the rough analog to the “runs to stills” terminology used in oil refining. The natural gas stream remaining after HGL extraction will then become an input to the “marketed production” projections in the natural gas supply module. The volumetric difference between estimated production (also known as “Gross Withdrawals” in EIA models) and inlets to processing plants is assumed to be included in the natural gas supply module and is not considered to be part of the HGL Supply projection module.

These ‘inlet’ volume projections are the volumes to be compared to the available capacity in order to assess the need for future investments in facilities in the sub-region. The recommended approach for projecting inlet volumes is to first eliminate gas that is considered to be ready for pipeline shipment without HGL extraction. In the United States, these streams are generally considered to be gas with wellhead Btu content less than 1000 Btu/Scf, and coalbed methane, which usually consists of methane and water. Overseas, a higher Btu cutoff point may be considered in warmer climates where the throughput volumes shown in data sources do not indicate a high percentage of HGL extraction.

Gas production having a Btu/scf content between 1000-1100 is marginal from a standpoint of requiring processing. This gas is often called “lean” gas in the United States. Whether HGL will be recovered from these streams is largely dependent on the outlook for HGL values, which will be exogenous inputs to the models from demand estimates. When capacity constraints exist in a field or basin, these volumes would be excluded first from the analysis, and added to the projection models only when external market signals are received that indicate capacity will be added to handle them. Where sufficient capacity already exists to accommodate these streams,

they can be included in the inlet volume estimates to produce a ‘potential’ HGL supply case run.

Gas from gas condensate wells, and gas from wells which are classified as oil wells typically contain enough HGLs to require processing to some extent, and these volumes should be included in the inlet estimates for the supply runs. They are considered the “must process” volumes in the model.

Gas Plant HGLs – Yields Projection Input Data Files

The projected wellhead production estimates, adjusted for gas quality considerations and capacity constraints, are used to create the input table for use in estimating HGL yields from gas processing facilities. Table 3 is an example of an input table that can be used with a yields estimator to project HGL yields from gas processing. The location is a US plant in this case, but this structure can be used for other regions.

**Table 3. HGL Supply from Gas Processing
Example Input File for use with HGL yields estimators**

Facility Type	Logistics Location Geocode	Facility Analog Model	Stream Type_Code	Pipeline_Btu_Limit	Volume Units	Volume_Period1	Volume_PeriodN
Gas Plant	Group 100	Opal	XXX	1035	Bcf/day	1,000,000	1,200,000

Table 3. Field Descriptor Definitions:

"Facility Analog Model" - a name or code that links to the data record that contains the “sub region analog” facility information that will be used by the yields model to estimate the supply for the sub-region. An example of an analog name would be to use the name “Algeria LNG” to point to the record that contains a summary of the NGL extraction capacity and configuration information for all the LNG facilities in Algeria. This capacity and configuration detail would then become a ‘mega plant’ that would be used to model the supply projections for the region.

“Stream Type Code” – designates the record that contains the quality and/or composition assumptions to be used for modelling yields. Depending on the amount of detail available and/or preferred, these records can be as complex as a complete assay for crude oil or a complete chromatographic analysis for natural gas, or as simply as the gravity or name designation for crude (for example, Bonny Light), or natural gas (“oil well gas”).

“Pipeline Btu_Limit” – specifies the maximum Btu/scf of the natural gas that can be delivered into the pipeline grid in the sub-region. This can be stored in the Logistics module along with the rest of the pipeline tariff and capacity information, or included in the HGL Yields input files depending on the approach to yields modelling that is chosen. This will serve as an upper limit on the amount of ethane that can be left in the natural gas stream as opposed to being sold as HGL.

“Volume_Period1...Volume_PeriodN” - these are the gas inputs for each time period for which a projection is being done.

Resource Representation – Extraction Process Supply Channel – Yields from Gas Plants

The final step in the supply projection workflow is the estimation of process yields. Given the desire to minimize costs, a sophisticated engineering model of yields is not essential. For HGLs, a commercial estimation tool such as the NGL Calculator developed by Midstream Energy Group can be used to calculate HGL yields given the information contained in the Supply Model Input File. In regions where ethane is produced as a petrochemical feedstock, a value for ethane and for natural gas are required in order to assess the marginal revenue extracting additional ethane over and above the minimum level required to meet pipeline Btu constraints.

The outputs from this model are the HGL from gas processing supply projections, which are stored in the HGL Supply Projection Database (See Figure 4.)

Resource Representation – Extraction Process Supply Channel – HGLs from Oil Refining

Projections for HGLs from oil refining also require an assessment of the throughput volumes for each refinery supply node. For oil, crude throughput in a sub-region may be a combination of domestic production and imports. These crude oil volumes should be generated by other modules which forecast oil supply and trade balances and are considered exogenous inputs to the HGL supply module. It is currently assumed that the HGL components from refining will be input into this module from other refinery sub-modules, but we have included the information that would be needed to generate these projections within the HGL module in this section.

HGLs from Oil Refineries – Crude Quality Input Data

If other supply modules such as the Liquid Fuels Module contain refinery models include estimated yields of HGLs, these volumes can be input directly into the HGL supply projection database as exogenous inputs. This design description assumes that the HGL module will be needed to produce these projections.

Table 4. is an example of an input table for use in estimating HGL yields from oil refining.

Table 4. HGL Supply from Oil Refining
Example Input File Structure for use with yields estimators

Facility Type	Logistics Location Geocode	Supply Node Analog Model	Node Capacity	Crude Type	Crude Throughput Period1	Crude Throughput PeriodN
Oil Refinery		Singapore	1.344 million bpd	Arab Medium		

If these volumes are not going to be brought in from another module in the Global Hydrocarbon Supply suite, an engineering consultant can create a high level yields model for each group of refinery assets to provide a “factored” estimate of HGL yields (and external HGL demand) for use in estimating refinery supply and demand for HGLs. This is basically a set of percentages that can be applied to the crude volumes and quality assumptions

to estimate the HGLs produced by the various refinery units, and to estimate the external volumes needed for gasoline blending or alkylation production.

These models will need to be reviewed and adjusted periodically if configurations are modified or crude types change significantly from those used to build the initial estimates, but they provide a way for non-technical personnel to create supply projections without extensive training in process simulation.

The HGL supply projections from the refinery nodal yield models are stored in the HGL Supply Production database. Demand projections should be passed to the overall demand module which accumulates total energy demand for use in determining net trade balances.

Resource Representation – Extraction Process Supply Channel – Supply Projection Database

This datafile contains the results of the supply projection modelling process. It is recommended that the projections be estimated on a component basis (ethane, propane, iso-butane, normal butane, pentanes+, ethylene, propylene, butylene) and stored at this level of detail. It is easier to combine 2 individual component volumes for a stream estimate to compare to reported historical data (for example, mixed butanes or refinery grade propane/propylene) than to separate a stream and attempt to balance trade flows without the details. Table 5. provides an example of fields to be included in this datafile

Table 5. HGL Supply Projection Database Fields - Example

Region ID	Level 1 Region Name	Logistics Location Geocode	Level 2 Resource Name	Supply Node Analog Model	Output Volume Units	Output_Type _HGL_C2	Output_Type_H GL_C3
2	MidEast_NA	Arab Gulf Jubail	Ghawar Field	Hawiyah Plant	MTA	100,000	28,000
2	MidEast_NA	Red Sea Yanbu		Western Saudi Arabia Center	MTA		500,000

Table 5. Field Descriptor Definitions:

“Output_Type_HGL_C2” is used to capture the projected yields of ethane

“Output_Type_HGL_C3” is used to capture the projected yields of propane

Additional output fields are added to incorporate the HGL component products for which projections are being made.

Product Types – Required for Interface with Demand Modules

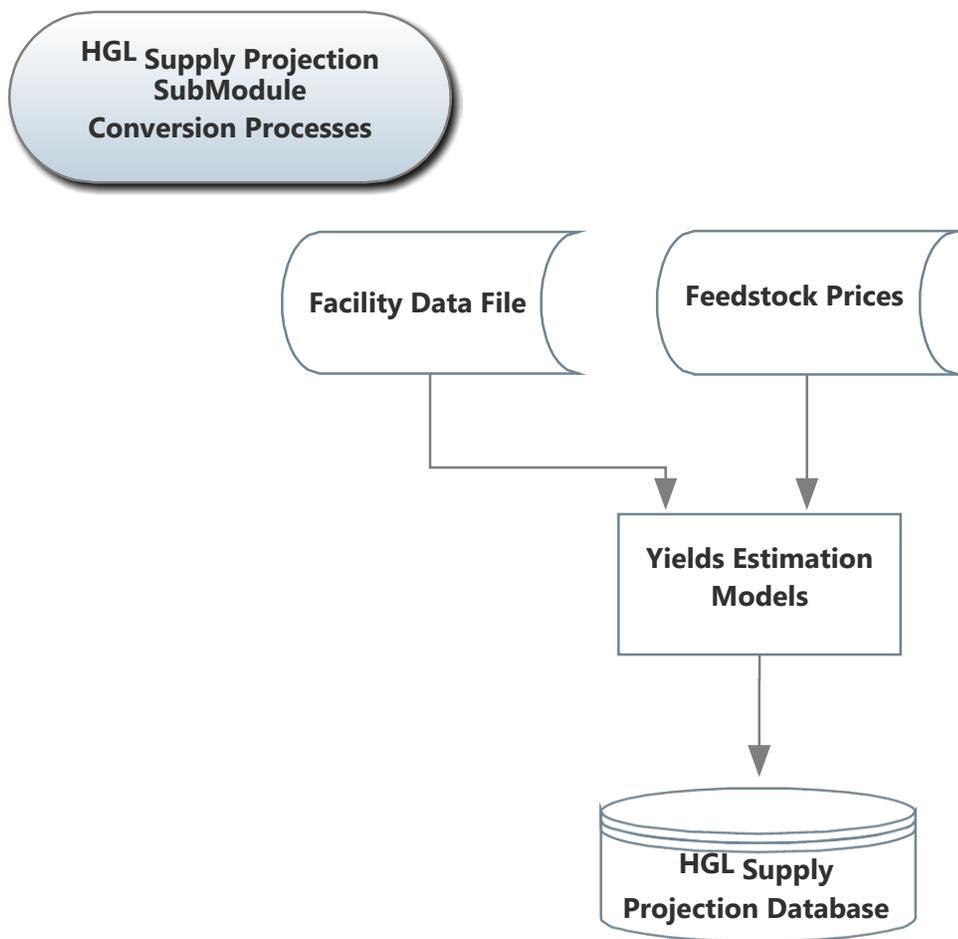
The current demand modules follow historical convention in providing estimated demand and pricing for “LPG”. Due to the increasing supply of propane projected to be supplied into the global petrochemical markets from the US over the next 15 years (going from net imports to over 1 million barrels per day by 2020), we recommend that the supply module capture yields for, at minimum, propane and mixed butanes for use in energy sector models. This would provide a level of detail that would allow volumes demanded as component products to be priced and directed to the appropriate global markets as they grow. The markets which are large enough to post prices for

HGLs typically post separate prices for propane and mixed butanes today. For HGL components forecast by this module which are used primarily for petrochemical production, ethane and natural gasoline (pentanes+), breaking down these volumes in the industrial demand portion of the demand models would provide a better way to capture changes in trade balances as ethane begins to displace naphtha as a feedstock in Eastern Canada and portions of Europe and Asia.

Resource Representation – Conversion Process Supply Channel – Work Flow

These products are primarily created from methane via methane reforming and include liquid fuels from coal, liquid fuels from natural gas (GTL), and liquid fuels from gas produced in biomass conversion (BTL). Figure 5. illustrates the workflow for determining HGL supply balances for liquid fuels produced by gas conversion processes.

Figure 5. Liquid Fuels from Gas Conversion Processes



Resource Representation – Conversion Process Supply Channel – Facility Data File

Like the Facility Data file used for the Extraction Process supply projections, the facility data file for use in producing supply projections for liquid fuels produced from gas via conversion processes contains capacity and configuration information. In addition, this facility file will need to contain an estimate of capital and operating costs in order to determine whether the cost of supplying these fuels into the market, including available

subsidies, exceeds their value when substituted for conventional hydrocarbon fuels.

These projects are usually initially supported by government research programs, supply agreements (such as contracts from the Department of Defense), or regulatory mandates. As such, there is usually a significant amount of information about them in the public domain.

There are currently a small number of these facilities, but like gas plants and refineries the technology involved tends to be relatively similar, so the nodal approach can be used for these projections as well. The initial nodes may only contain one facility, but the capacity and configuration can be increased as additional plants are added.

Resource Representation – Conversion Process Supply Channel – Feedstock Prices

This is an input series obtained from an external file, assumed to contain the common database of prices for various forms of energy and feedstock. These inputs include:

- Cost of coal delivered to the facility location
- Cost of biomass (wood pulp, crop residue, animal by-products)
- Cost of natural gas

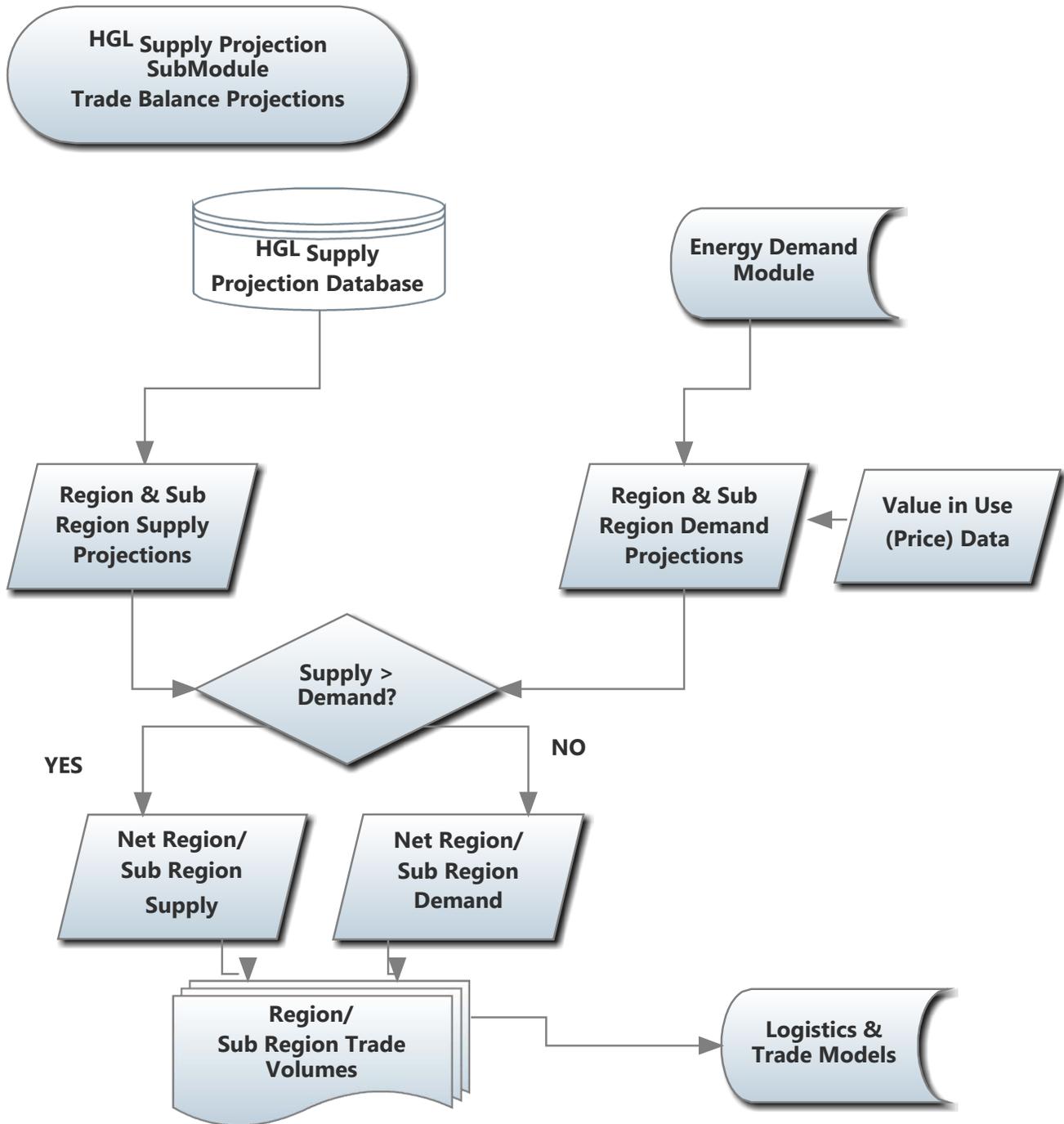
Resource Representation – Conversion Process Supply Channel – Yields Estimation Models

The process for developing supply projections for syngas to fuel conversion facilities is relatively straightforward. Using the information from the facilities datafile, a cost of supply is calculated. If the cost of production is greater than the value of the product net of subsidies, the projected supply volume would be set to -0-. If the cost of production is less than the value of the product, the supply volume can be set at the plant's production capacity adjusted for normal downtime and maintenance.

Resource Representation – Conversion Process Supply Channel – Net Regional Balances

Once the supply projections have been completed for a region and/or sub-region and stored in the HGL Supply Projection Database, the outputs can be used to determine the Net Regional Trade Balance for use in balancing markets outside the region. The workflow for this process is illustrated in Figure 6.

Figure 6. Net Regional Trade Balance Determination Process



The outputs from this calculation will be the preliminary trade balance that can be combined with logistics costs and price projections of value in use in other regions in order to determine projected trade flows over the forecast horizon.

In order to better align with the current demand models within the WEPS+ suite, and in particular the International Natural Gas Model, the minimum 7 supply regions can be expanded to the 16 currently included in the INGM model to create a net balance for the 16 regions. However, it will likely be difficult to determine

prices for all 16 regions. The current major price nodes for which price information is made available on a regular basis include:

- US Gulf Coast
- US MidContinent
- Western Canada
- Eastern Canada
- Northwest Europe
- South China
- East China
- Japan
- “Other Far East” (FEI or Far East Index) locations
- Saudi Contract Price (CP)
- Mediterranean Basin (“Med”)

Mexico, Central, Latin, and South America are notably missing from this list. There are no locations there where there is sufficient storage and an active enough spot market to provide transparency around pricing in this area. Assumed value in use estimates can be derived for these regions by determining the delivered cost of supply from the US Gulf Coast, capped at the value of product in Northwest Europe, plus freight, or the value of material from Saudi Arabia, plus freight. Other locations included in the 16 that are currently used in the INGM would require similar treatment. Use of this methodology would allow calculations of prices for the number of regions desired in the model.

Model Design Recommendations

Use Hydrocarbon Resources as the Initial Point Source for Supply Forecasting (Extraction Process)

The use of "resources" – defined as underground hydrocarbon reserves (for extraction process supply projections) and “facilities” as starting points to source initial data for supply modeling provides a way to create nodal forecasts that can be summed at various levels, mapped against geopolitical boundaries with geocoding, and linked to logistics network modules which describe physical supply aggregation points for storage and distribution in order to create estimates of delivered costs of supply to end use markets via tariff and freight data contained in a logistics module. The use of simulation as opposed to trend analysis using econometric tools is a departure from some suggested approaches, but provides a higher level of reliability in that variance analysis and calibration is easier in terms of identifying which portion of the model or dataset should be tweaked to produce a more accurate result.

The inventory of hydrocarbon subsurface assets (resources) around the world is large, but the number of resource regions (fields and basins) with expected production volumes large enough to be included in the analysis is much smaller. If necessary, the total volume available to be processed (gas plant inlets) can also be estimated for HGL forecasts for international locations without requiring detailed information about gas quality since these flows are usually tracked and reported at the project level. Most hydrocarbon resource development investments outside the United States and Canada are generally negotiated and sanctioned by host governments who publicize them as economic growth engines. Large facilities sometimes even have their own websites.

In terms of the size of the surface facilities universe, the Oil and Gas Journal's Annual Worldwide Gas Processing Survey³ lists a total of 1,983 gas processing plants in the entire world as of year-end 2013, with 82% of these located in the United States (649) and Canada (973) where information is readily available either via EIA's plant surveys or inexpensive industry data sources such as the Oil and Gas Journal document.

Modeling Considerations in Determining Node Composition

While the basic unit of sourcing for supply forecasting is recommended to be at the facility level, for efficiency in estimating supply for sub-regions and regions and capturing trade flows, the forecast volumes should be modelled as nodes. The definition of the attributes of the nodes has the greatest influence on the overall level of accuracy of the forecast, and on the ability of those who maintain and calibrate the model to do so efficiently and cost effectively. For instance, if a sub-region has significant refinery capacity concentrated along 2 coastal regions, with 2 separate ports handling logistics, it is generally best to create 2 separate facilities nodes to make it easier to validate trade flows. The use of logistics centers as aggregation points to tie individual facilities to offtake assets within sub-regions makes it easier to identify these natural nodal clusters.

An overly broad definition of a node, especially one defined using geopolitical boundaries as node descriptors, could result in significant rework when boundaries shift, regulatory policies change, and/or it becomes necessary to revise model assumptions in order to determine how to improve the accuracy of a forecast which is seen to drift versus historical actual results. Hence the addition of sub-regions for allocating hydrocarbon production between geopolitical entities, with the facility nodes also associated at the sub-region level.

In cases where reservoirs cross political boundaries, this concept would create gas plant HGL supply nodes which are asset clusters in distinct producing basins, some of which cross state or national boundaries. For instance, the extraction plants handling gas production from the Green River basin in Utah and Southern Wyoming would be tagged as a single supply node, with production aligned with a local market area that lies within 'economic reach' of a truck or railcar based on transportation costs. Surplus volumes for the node are then aligned with the distribution point that moves them to another region.

Detailed Data Capture Adds Analytical Flexibility

This principle of detailed data capture pervades the proposed design approach. When the initial NEMS models were developed, analysts had the benefit of very rapid computational power, but the disadvantage of limited availability of personnel with the programming skills to take on the design work involved in building, maintaining, and adapting a complex database structure to a changing market, much less the resources to populate databases with information that was likely in hard copy or microfilm format, and scattered across the country if available at all. The availability of 21st century rapid system prototyping tools and decades of regulatory filings, industry websites, trade association statistics, online records repositories, geolocation services such as Google Earth, and robot internet search capabilities can make the "set up" and maintenance tasks associated with updating

³ "New, expanded capacities return US, Canada to global lead", *Oil and Gas Journal*, June 2, 2014, online edition.

facilities information much simpler. Most systems designers prefer to focus on the intellectually stimulating portion of the solver programs, but attention to detail in initially setting up the datasets will pay off later on in increased accuracy, efficiency, and maintainability.

Economic Constraints – Cost of Supply

With the exception of the ethane recovery forecast, the HGL from extraction model does not calculate cost of supply or use it as a variable in forecasting production. HGL is considered a by-product of gas and oil extraction, and the volumes are a function of the quantity of HGL contained in the gas and the gas pipeline limitations around Btu and hydrocarbon dewpoint. Therefore, recoveries of HGL other than ethane are assumed to occur regardless of value in use if the gas is determined to require processing. The variable used to allocate HGL volumes to end users is value in use as opposed to cost, since the producer must accept a loss (incur a cost) if the cost of transporting the HGL to a market is higher than its value to the customer. The value in use is assumed to be provided by a price data file exogenous to the HGL module, since it is dependent on the alternative fuels available within the defined markets, the availability and cost of which is also determined by other modules.

Complexity of Programming

For projections of HGL supplied from extraction processes and crude oil refining, the level of accuracy needed is not great enough to require a process simulation package. A simple set of ‘factors’ for refinery clusters (nodes) and a simple commercial calculator for gas processing plants should be enough to produce component volume projections for use in these models.

The programming tools needed to produce HGL yields forecasts for conversion technologies, such as those used in CTL, GTL, or “BTL” can be as sophisticated as the models used in design simulations, or as straightforward as a simple high level yields model that relies on input assumptions around throughput volume, capacity, and price, which we are assuming in this design recommendation. The HGL module will produce projections of supply available at points where logistics facilities are available, allowing the calculation of the cost to move these volumes from the point of production to the point of distribution to regional markets. For projections of supply from conversion facilities, facility capital and operating expenses, together with the cost of their primary feedstocks, can be used to generate a delivered cost of supply from these nodes.

Calibration and Validation – Initial and Subsequent Model Runs – Region and Sub-Region Results

The use of a nodal model allows the results of the modelling process to be validated using historical data. Production statistics are often available at a facility level, especially for large facilities with a material impact on energy supply within a sub-region or region. In the US, facility operators submit data that can be aggregated the same way as the facility nodes in the model to test the projections and pinpoint areas where adjustments should be made.

Calibration and Validation – Initial and Subsequent Model Runs – Global Balances

The single best way to check the level of accuracy in modelling these balances is to compare the results of the projected trade balances with historical data from service providers who produce trade statistics. A number of companies track international and cross-border trade, and while not all major suppliers report their trade statistics

directly, many of their trading partners do. This makes it possible to develop reasonably accurate assessments of flows across and between regions. Having this data, which is broken down by product, allows the output of the supply model's trade balance results and the results of the logistics model solver which estimates the likely flow pattern to be compared to known results.

Data Requirements

A number of private research organizations, including Wood MacKenzie, offer specialty services that can provide datafiles to use in setting up the module. The EIA also has extensive datafiles on much of the information needed as a starting point in configuring the module. The data sources listed below are lower cost routes to acquiring additional data. The essential services recommended for subscription are some sort of price service, and a trade statistics service. These can be shared across module owners, as pricing and trade flows are needed for projections for many commodities.

Hydrocarbon Resource Database - Reserves and Production:

In addition to existing EIA data files, these resources are relatively low cost:

U.S. Geological Survey, Energy Resource Program
BP World Energy Survey
Shell (See 'projects' section of the website)
Exxon World Energy Survey
ENI World Oil and Gas Survey
Society of Petroleum Engineers
American Association of Petroleum Geologists
International Energy Association (IEA)
Natural Gas – Cedigaz - annual resource survey:
<http://www.cedigaz.org/products/annual-surveys-2013.aspx>

NOC website examples:

Iran - <http://www.nioc.ir/Portal/Home>
Oman – <http://www.oman-oil.com>

Facilities Information – Projects, Construction:

Low Cost –

Oil and Gas Journal Surveys - \$500-\$600 each in Excel format:

Worldwide Gas Processing Survey

Worldwide Refining Survey

Worldwide Gas Processing Construction Projects – includes planned construction for gas processing plants, LNG plants, and gas to liquids facilities

Higher Cost – International Projects:

Energy Intelligence publications, *Nefte Compass* (Russia), *Petroleum Intelligence Weekly*, *World Gas Intelligence*

Trade Statistics – cross border and international trade – all products:

GTIS – broadest geographical coverage

Datamyne – can provide custom reports, fewer countries than GTIS

PIERS – excellent coverage, fewer countries, highest price

Trade Statistics – HGL trade (propane, butanes)

Waterborne LPG Report (IHS) - historical trade balances, waterborne freight and common voyage arbitrage economics:

Logistics Costs and Vessel Statistics:

Fearnley's – weekly update on shipping costs

Clarksons – vessel fleet statistics, freight costs

Lloyd's Intelligence – vessel tracking, fleet statistics

Prices – HGLs

Argus International LPG (www.argus.media.com)

Oil Price Information Service (www.opis.net)

Yields Modelling – Refinery

Jacobs Consultancy, Vince DiVita

Yields Modelling – NGLs

Midstream Energy Group has a prototype calculator that can be adapted for use with multi-period files. Stan Keller is the demonstration contact at stan.keller@bizplace.com

Software Considerations

The HGL Supply module is designed to be a Business Intelligence System. These use databases, data normalization processes, a nodal approach, and intermediate processing to derive results designed to be used for business planning and strategy development. A number of vendors, including Oracle, SAP, and Microsoft can supply a 'template' system that avoids the cost of custom programming. The intermediate calculations, including the yields models, can be performed in Excel, or in the case of calculations involving decisions around product movements, an off the shelf linear programming package.

Uncertainty and Limitations

The goal of the HGL supply module is to produce HGL supply projections that are within 15% of the actual volumes appearing in the market. Given the limitation on the availability of detailed information on these specialized components, it will be difficult to directly measure these results, so a proxy goal would be to predict within 15% the overall volumes of HGL's in global trade as captured by the trade reporting services.

The use of facility nodal models instead of trend analysis should allow analysts to better calibrate results, as a considerable amount of detail is available around the output of large projects, which allows the derivation of estimates around sub-region balances even if country level data is not readily available, which is the case for many non-OECD countries.

References – Definitions

Hydrocarbon Gas Liquids – Products to be included in the projections from this module, based on United States specifications for Mont Belvieu, Texas:

“Saturated” Streams (non-olefinic):

Ethane (C2) – a stream typically containing no less than 95% of ethane and no more than 3% methane. This product would be supplied to industrial markets as petrochemical feedstock in global models.

Propane (C3) – a stream typically containing no less than 90% propane, having a maximum Reid Vapor Pressure of 208# psi

Iso-Butane (IC4) – an isomer of normal butane, a stream having no less than 96% isobutane and 3% propane

Normal Butane – a stream having no less than 94% normal butane

Mixed (or “Field Grade” Butanes” – this is the product code that would likely be used for supply outside the United States, as the combination of normal and isobutane. Many locations sell these as a stream, which can be fractionated at the delivery point or used in residential applications

Pentanes “Plus”, also known as Natural Gasoline or C5+ in the United States – a stream containing no less than 97% pentanes and heavier molecules, no more than 3% butanes, and having a Reid Vapor pressure of no more than 14# psi. This product would be supplied to industrial markets as petrochemical feedstock in global models.

“Unsaturated” Streams (olefinic):

Ethylene (C2=) – a stream containing 99% ethylene

Propylene (C3=) – a stream containing no less than 65% propylene (refinery grade)
