3-2014

Strategic Planning, Design and Development of the Shale Gas Supply Chain Network

Diego C. Cafaro
Universidad Nacional de Litoral

Ignacio E. Grossmann
Carnegie Mellon University, grossmann@cmu.edu

Published In
AIChE Journal, 60, 6, 2122-2142.

Strategic Planning, Design and Development of the Shale Gas Supply Chain Network

Diego C. Cafaro[1] *and Ignacio E. Grossmann*[2*]

[1] INTEC (UNL – CONICET), Güemes 3450, 3000 Santa Fe, ARGENTINA

[2] Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, U.S.A.

ABSTRACT

The long-term planning of the shale gas supply chain is a relevant problem that has not been addressed before in the literature. This paper presents a mixed-integer nonlinear programming (MINLP) model to optimally determine the number of wells to drill at every location, the size of gas processing plants, the section and length of pipelines for gathering raw gas and delivering processed gas and by-products, the power of gas compressors, and the amount of freshwater required from reservoirs for drilling and hydraulic fracturing so as to maximize the economics of the project. Since the proposed model is a large-scale non-convex MINLP, we develop a decomposition approach based on successively refining a piecewise linear approximation of the objective function. Results on realistic instances show the importance of heavier hydrocarbons to the economics of the project, as well as the optimal usage of the infrastructure by properly planning the drilling strategy.

KEYWORDS

Shale gas, supply chain, strategic planning, MINLP, solution algorithm

---

* Corresponding author. Tel.: +1 412 268 2230; fax: +1 412 268 7139.

E-mail address: grossmann@cmu.edu (I.E. Grossmann).

1. INTRODUCTION

Natural gas is the cleanest-burning fossil fuel. Natural gas extracted from dense shale rock formations has become the fastest-growing fuel in the U.S. and has the potential of becoming a significant new global energy source. Over the past decade, the combination of horizontal drilling and hydraulic fracturing has allowed access to large volumes of shale gas that were previously uneconomical to produce. The production of natural gas from shale formations has reinvigorated the natural gas and chemical industries in the U.S. The Energy Information Administration projects U.S. shale gas production to grow from 23% to almost 50% of the total gas production in the next 25 years.[1] Shale gas is found in plays containing significant accumulations of natural gas, sharing similar geologic and geographic properties. A decade of production has come from the Barnett Shale play in Texas. Experience gained from Barnett Shale has improved the efficiency of shale gas development around the country. Today, one of the most productive plays is the Marcellus Shale in the eastern U.S., mainly in Pennsylvania. Regarding both economic and environmental impacts, the long-term planning and development of the shale gas supply chain network around each play is a very relevant problem. However, to the best of our knowledge, it has not been addressed before in the literature.

The raw gas extracted from shale formations is transported from wellbores to processing plants through pipelines. The processing of shale gas consists of the separation of the various hydrocarbons and fluids from the pure gas (methane) to produce what is known as “pipeline quality” dry natural gas.[2] This means that before the natural gas can be transported by midstream distributors, it must be purified to meet the requirement for pipeline, industrial and commercial uses. The associated hydrocarbons (ethane, propane, butane, pentanes and natural gasoline) known as Natural Gas Liquids (NGLs), are valuable byproducts after the natural gas has been purified and fractionated. These NGLs are sold separately (usually through dedicated pipelines) and have a variety of different uses, including enhancing oil recovery in wells, providing raw materials for oil refineries or petrochemical plants, and as sources of energy.[3] One of the most critical issues in the design and planning of the shale gas supply

chain network is the sizing and location of new shale gas processing and fractionation plants (as well as future expansions) due to their high cost.

On the other hand, the number of wells drilled in each location can dramatically influence costs and the ecological footprint of natural gas operations.[4] The ability to drill multiple wells from a single location (or "pad") is seen as a major technological breakthrough driving natural gas development as for instance has happened in the Marcellus Shale. The utilization of multi-well pads also has large environmental and socio-economic implications given that as many as 20 or more natural gas wells and associated pipeline infrastructure can be concentrated in a single location. Furthermore, the total amount of industrial activity can be compressed as these wells can be drilled in rapid-succession and the technology now exists to perform hydraulic fracturing stimulations on multiple wells simultaneously. Hence, another key decision tackled by this paper is the drilling strategy, i.e. how many wells to set up or add on existing well pads at every time period.

Another critical aspect in the shale gas production is water management. Shale gas production is a highly water-intensive process, with a typical well requiring around 5 million gallons of water normally over a 3-month period to drill and fracture, depending on the basin and geological formation.[5] The vast majority of this water is used during the fracturing process, with large volumes of water pumped into the well with sand and chemicals to facilitate the extraction of the gas. Although increasing amounts of water are being recycled and reused, freshwater is still required in high quantities for the drilling operations as flowback water usually only represents about 25-30% of the water injected into the well. The need for freshwater is an issue of growing importance, especially in waterscarce regions and in areas with high cumulative demand for water, leading to pressure on sources and competition for water withdrawal permits. Therefore, a long-term planning model for the development of shale gas fields should also account for water availability.

The goal of this paper is to develop a mixed-integer nonlinear programming (MINLP) model for the sustainable long-term planning and development of shale gas supply chains, which should optimally determine: (a) the number of wells to drill on new/existing pads; (b) the size and location of new gas

processing plants (as well as future expansions); (c) the section, length and location of new pipelines for gathering raw gas, delivering dry gas, and moving NGLs; (d) the location and power of new gas compressors to be installed, according to the flowrate at every line; and (e) the amount of freshwater coming from available reservoirs used for well drilling and fracturing, so as to maximize the economic results (NPV-based approach) over a planning horizon comprising 10 years.

Literature Review

Some of the first papers in the optimal design and planning of supply chains were published in the literature 40 years ago.[6] A complete review on more recent developments in supply chain optimization problems can be found in the work of Melo et al.[7] Regarding the strategic planning of natural gas supply chains, Durán and Grossmann[8] propose a superstructure representation, an MINLP model and a solution strategy for the optimal synthesis of gas pipelines, deciding on the gathering pipeline system configuration, compressors power and pipeline pressures. Iyer et al.[9] propose a multiperiod MILP model for the optimal planning and scheduling of offshore oilfield infrastructure investment and operations. Since the resulting model becomes intractable due to the large-scale, nonlinear reservoir equations are approximated through piecewise linear functions. Van den Heever and Grossmann[10] propose a multiperiod generalized nonlinear disjunctive programming model for oilfield infrastructure planning, whose optimal solution is found through a bilevel decomposition method. In this model, the number of wells is given beforehand through a fixed drilling plan. More recently, Gupta and Grossmann[11] address some new features of the same problem, accounting for all three components (oil, water, and gas) explicitly in the formulation. They also incorporate more accurate estimations of the nonlinear reservoir behaviour, variable number of wells for each field (to capture drill rig limitations) and facility expansions, including their lead times.

On the other hand, some work has also been reported on the optimization of the operation of shale gas fields. Rahman et al.[12] present an integrated optimization model for hydraulic fracturing design, accounting for fracture geometry, material balances, operational limitations, characteristics of the gas

formation, and production profiles. By combining genetic algorithms and evolutionary techniques, improved hydraulic fracturing designs reduce the treatment (stimulation) costs up to 44% at the expense of a 12% reduction in the gas production. Knudsen et al.[13] propose a Lagrangean relaxation approach for scheduling shut-ins times in tight formation multi-well pads, so as to stimulate the shale gas production in different wells to comply with the gas rates required by the distribution company. In that work, a proxy model captures the physics during shut-ins operations. Based on the proxy model results, the time domain is discretized into daily time periods, and an MILP model is then solved using Lagrangean relaxation techniques.

Finally, few recent publications deal with the strategic and operational management of water resources and other environmental concerns in the development of shale gas plays. Mauter et al.[14] argue that strategic planning by both companies and regulatory agencies is critical to mitigate the environmental impacts of unconventional extraction. Rahm and Riha[15] attempt to determine water resource impacts of shale gas extraction, from regional, collective based perspectives, seeking to balance the need for development with environmental concerns and regulatory constraints. Yang and Grossmann[16] present an MILP formulation whose main objective is to schedule the drilling and fracturing of well pads to minimize the transportation, treatment, and freshwater acquisition costs, as well as treatment infrastructure, while maximizing the number of well stages to be completed within the time frame. The goal is to find an optimal short-term fracturing schedule, the water recycling ratio, and the need for additional impoundment and treatment capacity.

Like most enterprise-wide optimization (EWO) problems, the strategic planning of the shale gas supply chain has great economic potential. Considerable effort has been spent towards the solution of EWO problems during the last 20 years, particularly in the field of oil and gas production.[17] But none of them has been focused on the shale gas supply chain. The shale gas production has its own peculiarities, and is a problem of very recent development.[18] In fact, one of the major barriers is the size and complexity of computational optimization models for achieving the goal of EWO.[19] The strategic planning of shale gas infrastructure consists on the design of large supply chains, including

well-pads, processing plants, compressors, product delivery nodes, and the complex pipeline network transporting shale gas and the resulting hydrocarbons. As concluded by Oliveira et al.,[20] careful evaluation of the investment options in this kind of problems has particular importance, and the use of efficient decision-making tools that capture the problem complexity becomes crucial.

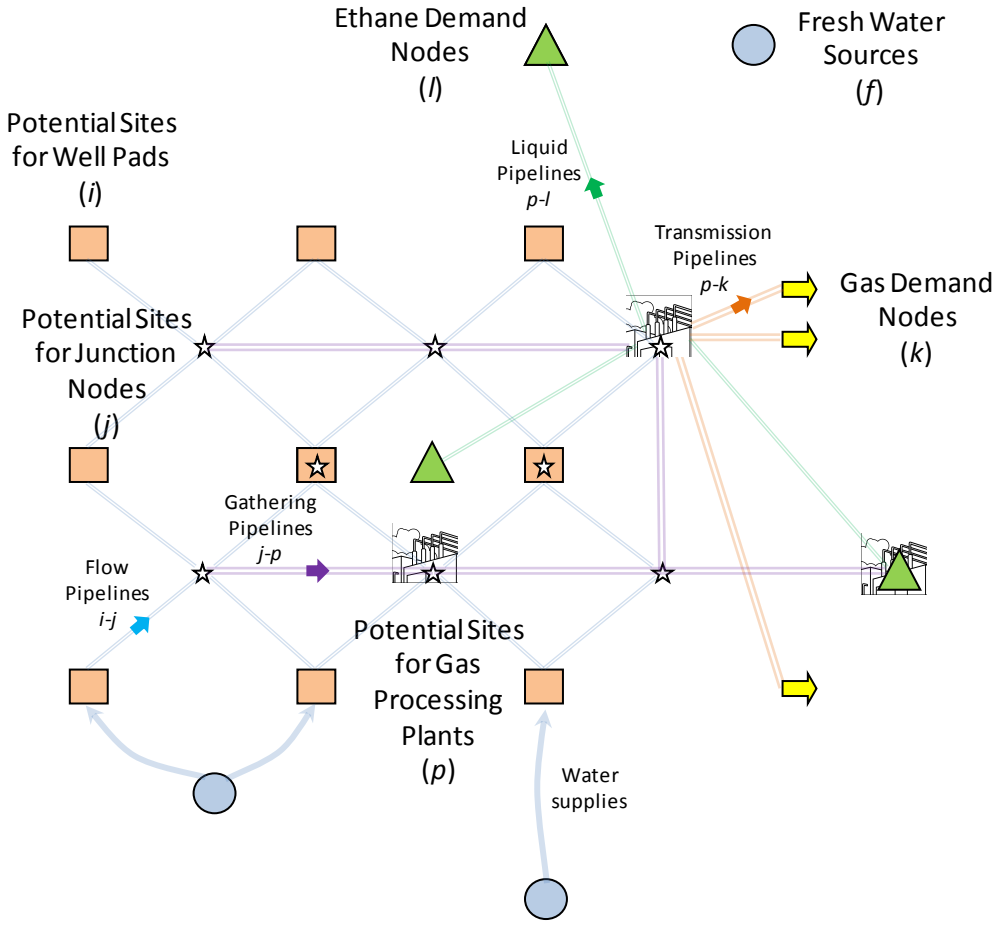

Figure 1. A simplified superstructure of the shale gas supply chain (for the sake of clarity, only few arcs of each type are drawn).

2. PROBLEM DESCRIPTION

We address the problem of determining the optimal design for a shale gas supply chain network, the well drilling and hydraulic fracturing strategy over the planning horizon, together with the size and location of gas separation plants, compressors and pipeline infrastructure, in order to maximize the net present value of the project. This problem can be formally stated as follows.

A comprehensive shale gas supply chain network superstructure like the one depicted in Figure 1 is given. It includes: (a) potential or existing well pads where new wells can be drilled and hydraulically fractured over the planning horizon (nodes $i \in I$), (b) potential or existing junction nodes where shale gas flows coming from nearby well pads converge (nodes $j \in J$), (c) potential or existing flow pipelines connecting nodes i and j, (d) candidate sites for the installation/expansion of a new/existing shale gas processing plants (nodes $p \in P$), (e) potential/existing gathering pipelines connecting junction nodes j with plant sites p, (f) demand nodes for dry natural gas (nodes $k \in K$) and ethane (nodes $l \in L$), (g) potential/existing transmission and liquid pipelines connecting plant sites p with nodes k and l, respectively, and (h) freshwater source nodes from where the water required for drilling and fracturing new wells can be supplied.

A strategic long term planning horizon is considered. In this paper, a planning horizon of 10 years is considered, and is divided into 40 time periods (quarters). The reasons for this time discretization is as follows: (1) Gas prices normally exhibit a seasonal behavior with a high peak in the winter. (2) The drilling and completion of wells normally takes between 50 and 90 days, plus the following 20 days during which the well does not produce a steady stream of gas, but a flowback of water that is captured and stored for further treatment. Overall, approximately 90 days (three months) are required since the well-pad is set up and wells start to be drilled until they begin to produce a steady flow of shale gas. (3) Freshwater availability in some water-scarce regions is strongly seasonal, and can be a critical issue if high cumulative demand for water leads to pressure on sources and competition for water withdrawal permits.

Besides the network superstructure and the time horizon, the productivity profile of every well at any location is assumed to be deterministic and known beforehand. Dry and semi-dry shale gas wells exhibit many of the same characteristics: an early peak in the gas rate from the sudden release of gas stored in pores and natural fracture networks, followed by a long transient decline in the production rate. Such decline in the rate is caused both by pressure loss and the inherently low permeability of shale rocks. In this problem, the well productivity (measured in Mm^3/day) is represented by a piecewise constant

function of the well age. The parameter $pw_{i,\tau}$ stands for the production rate of a shale gas well of age τ (given in quarters) drilled in location i (see Figure 2). Moreover, the shale gas composition, and particularly its "wetness" (% of hydrocarbons others than methane), are assumed to be known and independent of both the well site and its age. This assumption can be relaxed as will be discussed later in this paper.

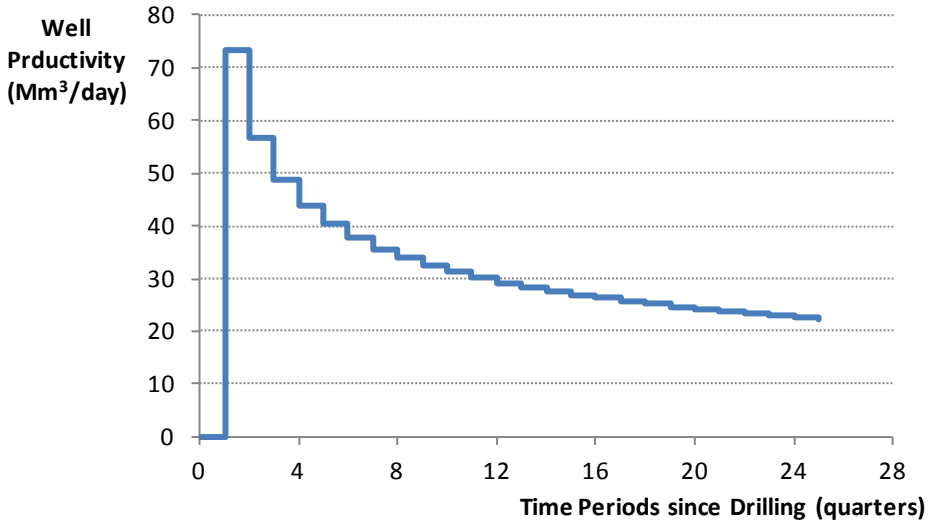

Figure 2. Piecewise constant well productivity profile.

Regarding the pipeline infrastructure, gas and liquid pipelines must be considered separately. On the one hand, gas pipelines (transporting either raw or processed gas) are assumed to handle an ideal mixture of ideal gases. Raw gas pipelines connecting nodes i to j (well pads to junction nodes) and j to p (junction nodes to plants) operate at medium-low pressures, while transmission pipelines p-k supplying gas demand nodes from processing plants operate at higher pressures. For simplicity, gas suction/discharge pressures at every node of the network are assumed to be given constant values. These are as follows: (i) shale gas discharge pressure at the well pads is Pd_i, (ii) junction nodes receive the shale gas at a pressure of $Ps_j < Pd_i$, (iii) compressor stations installed at junction nodes increase the pressure from Ps_j to Pd_j, to make the gas flow towards processing plants, (iv) the shale gas pressure at the inlet of processing plants is $Pi_p < Pd_j$, (v) processing plants deliver dry gas at a pressure of Po_p, (vi) compressor stations installed at the outlet of processing plants increase the dry gas pressure from $Po_p =$

Ps_p to Pd_p before sending flows to markets, and (vii) gas demand nodes receive dry gas at a pressure of $Pr_k < Pd_p$. By fixing such values, the maximum flow of a gas pipeline is directly proportional to the pipeline diameter raised to the power of 2.667, and the proportionality factor depends on the gas properties, the input/output pressures and the pipeline length.[8][21] Moreover, compressors are assumed to be adiabatic and their power is directly proportional to the gas flow, since the compression ratio is a given parameter. More details are given in the Appendix.

On the other hand, liquid pipelines transport hydrocarbons like ethane, propane, butane, pentanes and natural gasoline (NGLs) in liquid state from separation plants to either petrochemical plants or LPG (liquefied petroleum gases) distribution facilities. In this problem, all NGLs except ethane are assumed to be separately sold to customers near the processing plants, while ethane is continuously delivered to petrochemical plants by dedicated pipelines. The maximum flow in liquid pipelines is assumed to be directly proportional to the pipeline section since a maximum mean velocity is imposed.

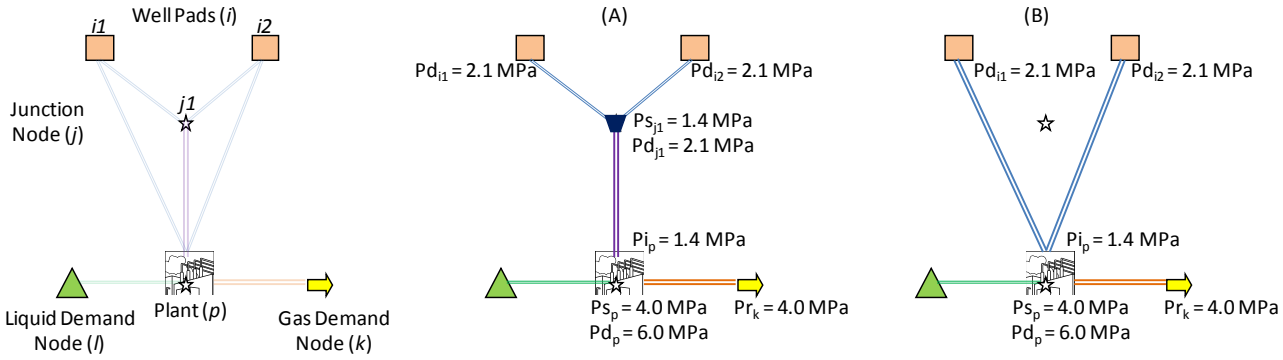

Figure 3. Simplified network superstructure and alternative network designs.

An illustrative example comparing two network designs is presented in Figure 3. The shale gas produced at two different well pads *i1* and *i2* is sent to a processing plant in two alternative ways: (A) through an intermediate junction node *j1*, or (B) directly, through separate lines. Typical values for the suction and discharge pressures at each node are given in the figure. The example also reveals one of the trade-offs to be determined by the model. Option (A) requires a compressor station at node *j1*, but pipelines are smaller in diameter and shorter than in option (B) which does not require a compressor.

Pipeline and compressor costs with regards to their size (usually determined by economies of scale functions) are the key to determine which option is the most convenient one.

Finally, freshwater consumption, mainly for hydraulic fracturing, is considered to be a fixed amount required during the drilling period, which depends on the well-pad location and the possibility of reusing the flowback water. The selection of optimal sources for water supply is a key model decision, but no details on the water transportation logistics are considered at this planning level. Other operational issues like flowback water capture, treatment and final disposal, as well as planning shut-ins and well stimulations are also out of the scope of this work.

Given all the items described above, the goal is to optimally determine: (a) the number of wells to drill on new/existing pads at every trimester; (b) the size and location of new gas processing plants (as well as future expansions); (c) the section, length and location of new pipelines for gathering raw gas, delivering dry gas, and transporting NGLs; (d) the location and power of new gas compressors to be installed, and (e) the amount of freshwater from available reservoirs for well drilling and fracturing so as to maximize the Net Present Value (NPV) of the project.

Assumptions

The main assumptions have already been discussed and can be summarized as follows:

(1) Shale gas is assumed to be an ideal mixture of ideal gases.

(2) The composition, and particularly the shale gas "wetness", are known constants independent of the well location. The relaxation of this assumption is discussed after the model presentation.

(3) The planning horizon is discretized in time periods, commonly quarters.

(4) Multiple wells can be drilled in a single pad over one time period, although not necessarily at the same time. It is assumed that all of them are hydraulic fractured and completed within the same time period they are drilled.

(5) Wells start to produce shale gas in the period following the drilling period. Once the wells are completed, their production cannot be delayed by shutting them in. This assumption can be relaxed as shown in the Appendix.

(6) After the well is completed, its productivity rate is a piecewise constant function in terms of the well age. In other words, a decreasing function as depicted in Figure 2 is assumed to be given.

(7) Multi-well pads can be set up, and multiple wells in the same pad can be drilled, fractured and completed during the same period. However, an upper bound is given due to technology limitations. Moreover, the total number of wells that can be drilled in the same pad over the given time horizon is also bounded.

(8) The pressure at pipelines transporting raw gas from well pads to junction nodes decreases from Pd_i to Ps_j as a function of their length[21] (for further details see the Appendix). The same is also valid for pipelines transporting raw gas from junction nodes to processing plants (from Pd_j to Pi_p), and dry gas from plants to demand nodes (from Pd_p to Pr_k).

(9) All the gas pressures (Pd_i at the outlet of pad i; Ps_j at the inlet of the junction node j; Pd_j at the outlet of j; Pi_p at the inlet of the plant p; $Po_p = Ps_p$ at the outlet of plant p; Pd_p at the outlet of the p-compressor station; and Pr_k at the gas demand node k) are given. Relaxing this assumption would imply solving a much more complex optimization problem.[8] Although pressure optimization is out of the scope for this model, it will be shown later that varying pressure levels within normal values does not lead to major changes in the optimal solution.

(10) The liquid pipeline flow is bounded by a maximum mean velocity (commonly, 1.5 m/s).

(11) Centrifugal pumps have negligible costs compared to processing plants, pipelines and gas compressors.

(12) Shale gas processing plants separate NGLs (namely ethane, propane, butane, pentanes and natural gasoline) from the shale gas (methane), also removing H_2S, CO_2, N_2 and H_2O; and finally delivering the methane to consumer markets. All NGLs except ethane are sold to nearby markets, while ethane is sent to chemical plants by dedicated pipelines.

(13) Concave cost functions of the form $f(x) = c\,x^r$ (with $0 < r < 1$ and $c > 0$) are assumed for: (a) the cost $f_a(x_a)$ of a shale gas processing plant with a capacity of x_a MMm^3/day,[22] (b) the cost $f_b(x_b)$ of a pipeline of diameter x_b, (c) the cost $f_c(x_c)$ of a compressor station of power x_c,[23] and (d) the cost $f_d(x_d)$ of drilling and hydraulically fracturing x_d wells during the same quarter year.

(14) Pipeline diameters are treated as continuous variables, but after the solution they are rounded up to the closest commercial diameter. A rigorous model would explicitly handle discrete size diameters, but this is out of scope of this work.

3. MATHEMATICAL FORMULATION

The optimization problem for the long-term planning, design and development of the shale gas supply chain is formulated in terms of a mixed-integer nonlinear programming (MINLP) model described in the following sections.

3.1 Model Constraints

The feasible region of the model is determined by a set of linear constraints. They are grouped into five blocks: Shale Gas Production; Flow Balances; Plants, Pipelines and Compressors Sizing; Plant, Pipelines and Compressors Costing; Water Supplies; and Maximum Demands.

3.1.1 Shale Gas Production

Number of Wells Drilled in a Pad. The number of wells drilled, fractured and completed in the multi-well pad i during the period t is represented by the variable $N_{i,t}$. Its value is determined by eq. (1) in terms of 0-1 variables $y_{i,t,n}$, one of which is equal to one to make $N_{i,t} = n$. The index n stands for an integer number greater or equal to zero and lesser or equal to $\overline{n}_i$, where $\overline{n}_i$ is the maximum number of wells that can be drilled during a single quarter in pad i. For the examples solved in the results section, the value of $\overline{n}_i$ varies from 2 to 4. Moreover, the total number of wells that can be drilled in a pad over the given planning horizon is bounded by eq. (3) to a maximum of $\overline{N}_i$. The current trend in shale gas production is to increase this number as much as possible to reduce the environmental impact.[4]

$$N_{i,t} = \sum_{n=0}^{\overline{n}_i} n\, y_{i,n,t} \qquad \forall i \in I, t \in T \tag{1}$$

$$\sum_{n=0}^{\overline{n}_i} y_{i,n,t} = 1 \qquad \forall i \in I, t \in T \tag{2}$$

$$\sum_{t \in T} N_{i,t} \le \overline{N}_i \qquad \forall i \in I \tag{3}$$

Shale Gas Production at Every Well-Pad. As stated in the model assumptions, the total production of shale gas (including methane, ethane and other NGLs) in a well-pad *i* at a certain period *t* depends on the age of every active well at that time. If $pw_{i,a}$ is a model parameter standing for the productivity (in Mm^3 of shale gas per day) of a well drilled in pad *i* , *a* quarters before the current time period *t*, then the total daily production coming from all the wells in pad *i* can be determined through eq. (4). Note that at time *t*, the age of a well drilled in time period $\tau < t$ is $a = t - \tau$. Moreover, wells of age "0" (being drilled and fractured) do not produce gas until the following period (see Figure 2).

$$\sum_{\tau=1}^{t-1} N_{i,\tau}\ pw_{i,t-\tau} = SP_{i,t} \qquad \forall i \in I, t > 1 \tag{4}$$

Methane, Ethane and other NGLs Produced at Well Pads. Since the shale gas composition at every well is assumed to be the same (uniform gas "wetness"), the production of such fuels within the shale gas stream coming from each well pad can be easily determined from eqs. (5), (6) and (7).

$$SP^G_{i,t} = gc\ SP_{i,t} \qquad \forall i \in I, t > 1 \tag{5}$$

$$SP^E_{i,t} = ec\ SP_{i,t} \qquad \forall i \in I, t > 1 \tag{6}$$

$$SP^L_{i,t} = lc\ SP_{i,t} \qquad \forall i \in I, t > 1 \tag{7}$$

where *gc* is the volume methane composition, *ec* is the ethane composition, and *lc* is the remaining hydrocarbons composition. If these parameters become dependent on the well location, the model

structure must be modified in order to preserve the linearity of model constraints. This will be discussed in a later section.

3.1.2 Flow Balances

Stream Flows from a Well Pad to Junction Nodes. Shale gas production at a certain pad during a time period is sent to one or more junction nodes (depending on the network design), which is controlled by eq. (8).

$$SP_{i,t} = \sum_{j \in J} FP_{i,j,t} \qquad \forall i \in I, t > 1 \tag{8}$$

The model variable $FP_{i,j,t}$ stands for the daily shale gas flowing from pad i to junction node j during period t. By simple extension of eqs. (5), (6) and (7), individual hydrocarbon flows in the shale gas stream ($FP^{G}_{i,j,t}$, $FP^{E}_{i,j,t}$, $FP^{L}_{i,j,t}$) can be easily obtained.

Flow Balances at Junction Nodes. Eq. (9) states that the sum of incoming shale gas flows at a certain junction node equals the sum of outgoing streams sent to one or more processing plants, depending on the network design. Under the given assumptions, flow splitting at well pads and junction nodes are both allowed.

$$\sum_{i \in I} FP_{i,j,t} = \sum_{p \in P} GP_{j,p,t} \qquad \forall j \in J, t > 1 \tag{9}$$

Similarly to variable $FP_{i,j,t}$, individual fuel flows can also be derived from the shale gas stream flowing between nodes j and p ($GP_{j,p,t}$, $GP^{G}_{j,p,t}$, $GP^{E}_{j,p,t}$, $GP^{L}_{j,p,t}$).

Flow Balances at Separation Plants. Assuming that all the methane from the shale gas flows processed at plant p is separated and sent to one or more dry gas demand nodes k, eq. (10) is added to the formulation. $TP_{p,k,t}$ is the flow of dry gas (methane) transported through pipeline p-k during period t.

$$\sum_{j \in J} GP^{G}{}_{j,p,t} = \sum_{k \in K} TP_{p,k,t} \qquad \forall p \in P, t > 1 \tag{10}$$

The same also applies for ethane flows, which are received with the shale gas, separated and pumped to one or more petrochemical plants *l* in liquid state through dedicated pipelines *p*-*l*, at a rate of $LP_{p,l,t}$ tons per day, during the entire period *t*.

$$s_g^E \sum_{j \in J} GP^E{}_{j,p,t} = \sum_{l \in L} LP_{p,l,t} \qquad \forall p \in P, t > 1 \tag{11}$$

s_g^E is the specific gravity of ethane in standard conditions, given in ton/MMm^3.

Finally, other NGLs from junction nodes are processed, and sold to nearby markets at a rate of $NP_{p,t}$ tons per day as stated by eq. (12).

$$s_g^L \sum_{j \in J} GP^L{}_{j,p,t} = NP_{p,t} \qquad \forall p \in P, t > 1 \tag{12}$$

3.1.3 Plants, Pipelines and Compressors Sizing

Separation Plants. The total processing capacity of a plant *p* at time *t* ($SepCap_{p,t}$) is given in MMm^3 of shale gas per day, and can be calculated from its capacity at the previous period ($t - 1$) plus the capacity expansion started at the beginning of period ($t - \varpi$), i.e. $SepInst_{p,t-\varpi}$. In other words, it is assumed that separation plants installations/expansions take ϖ time periods, as stated in eq. (13).

$$SepCap_{p,t} = SepCap_{p,t-1} + SepInst_{p,t-\varpi} \qquad \forall p \in P, t > 1 \tag{13}$$

Upper Bound on the Shale Gas Flows Converging to a Separation Plant. The sum of the shale gas flows coming from one or several junction nodes to a single separation plant during every period *t*, should not exceed its processing capacity as expressed by eq. (14).

$$\sum_{j \in J} GP_{j,p,t} \le SepCap_{p,t} \qquad \forall p \in P, t > 1 \tag{14}$$

Installation of Gas Pipelines. As shown in the Appendix, given the gas inlet and outlet pressures, the fluid properties and the pipeline length, maximum gas flows are directly proportional to the pipeline

diameter to the power of 2.667. It is also assumed that both raw and dry gases are ideal mixtures of ideal gases. In order to preserve linearity in the constraints, the diameter of the pipeline installed between a pair of nodes during a certain time period (a model decision) is substituted by a variable that stands for such diameter raised to the power of 2.667. In other words, the model variables $DFP_{i,j,t}$, $DGP_{j,p,t}$ and $DTP_{p,k,t}$ stand for the diameters of the pipelines installed at period t between nodes i-j, j-p, and p-k, respectively, raised to the power of 2.667.

In summary, the gas pipeline flows with regards to pipeline diameters are calculated from eqs. (15), (16) and (17).

$$FPFlow_{i,j,t} = k_{i,j}\ l_{i,j}^{-0.5} DFP_{i,j,t} \qquad \forall i \in I, j \in J, t > 1 \tag{15}$$

$$GPFlow_{j,p,t} = k_{j,p}\ l_{j,p}^{-0.5} DGP_{j,p,t} \qquad \forall j \in J, p \in P, t > 1 \tag{16}$$

$$TPFlow_{j,k,t} = k_{p,k}\ l_{j,k}^{-0.5} DTP_{j,k,t} \qquad \forall p \in P, k \in K, t > 1 \tag{17}$$

Due to assumption 9, parameters $k_{i,j}$, $k_{j,p}$ and $k_{p,k}$ take fixed values that can be calculated as shown in the Appendix. Distances between every pair of nodes ($l_{i,j}$, $l_{j,p}$ and $l_{p,k}$) are also given data.

Maximum Gas Flow between a Pair of Nodes. The maximum gas flow between every pair of nodes depends on the size of the pipelines installed in previous periods, plus the additional flow capacity added due to a recent pipeline construction, as stated in eqs. (18), (19) and (20). It is assumed that pipelines are installed from period ($t - \tau p$) to ($t - 1$) and are not able to transport gas until the period t, with τp being the pipeline construction lead time in quarters.

$$FPCap_{i,j,t} = FPCap_{i,j,t-1} + FPFlow_{i,j,t-\tau p} \qquad \forall i \in I, j \in J, t > 1 \tag{18}$$

$$GPCap_{g,j,t} = GPCap_{g,j,t-1} + GPFlow_{j,p,t-\tau p} \qquad \forall j \in J, p \in P, t > 1 \tag{19}$$

$$TPCap_{p,k,t} = TPCap_{p,k,t-1} + TPFlow_{p,k,t-\tau p} \qquad \forall p \in P, k \in K, t > 1 \tag{20}$$

Finally, shale gas and dry gas flows at every time period are bounded by the flow capacity connecting every pair of nodes, as enforced by eqs. (21), (22) and (23).

$$FP_{i,j,t} \leq FPCap_{i,j,t} \qquad \forall i \in I, j \in J, t > 1 \tag{21}$$

$$GP_{j,p,t} \leq GPCap_{j,p,t} \qquad \forall j \in J, p \in P, t > 1 \tag{22}$$

$$TP_{p,k,t} \leq TPCap_{p,k,t} \qquad \forall p \in P, k \in K, t > 1 \tag{23}$$

Installation of Liquid Pipelines. By assumption 10, a maximum mean velocity is imposed to liquid flows to make sure that head losses remain at specified values. In liquid pipeline network design, a typical value used is $v^{\max} = 1.5$ m/s. Under such an assumption, liquid flows are directly proportional to the pipeline section, and by extension, directly proportional to the pipeline diameter raised to the power of 2. As for gas pipelines, the diameter of a liquid pipeline installed between a gas processing plant p and a petrochemical plant l during a certain time period (a model decision) is substituted by an analogous variable, which stands for such diameter to the power of 2 (variable $DLP_{p,l,t}$). As a result, gas pipeline flows (given in tons per day) with regards to pipeline diameters are calculated by eq. (24).

$$LPFlow_{p,j,t} = k_{p,l}\ DLP_{p,l,t} \qquad \forall p \in P, l \in L, t > 1 \tag{24}$$

where $k_{p,l} = 3600 \cdot 24 \cdot \rho\ \pi\ v_{p,l}^{\max} / 4$, ρ is the liquid (ethane) density given in ton/m^3, and $v_{p,l}^{\max}$ the maximum mean velocity, in m/s.

Maximum Flow in Liquid Pipelines. Similarly to gas pipelines, the model decides when to install a new pipeline and its corresponding size. Eq. (25) determines the flow capacity of each liquid pipeline at every time period (in ton/day), while constraint (26) imposes such value as an upper bound on the liquid flow from p to l during period t.

$$LPCap_{p,l,t} = LPCap_{p,l,t-1} + LPFlow_{p,l,t-\tau p} \qquad \forall p \in P, l \in L, t > 1 \tag{25}$$

$$LP_{p,l,t} \leq LPCap_{p,l,t} \qquad \forall p \in P, l \in L, t > 1 \tag{26}$$

Power of Compressors. If the suction and discharge pressures are given (see assumption 9), and assuming that compressors are adiabatic, a simple expression can be derived in order to calculate the required compression power (in kW) as shown in the Appendix. Under such assumptions, the required power is directly proportional to the total flow of gas being compressed. In the case of raw gas, compressed at junction nodes and sent to processing plants, the total power installed up to time t ($JCP_{j,t}$) must be greater or equal than the power demanded by the total flows of raw gas compressed by j at time t, as expressed by eq. (27).

$$JCP_{j,t} \geq kc_j \sum_{p \in P} GP_{j,p,t} \qquad \forall j \in J, t > 1 \tag{27}$$

Similarly, the power of compressors installed at the outlet of the processing plant p up to time t ($PCP_{p,t}$) sending dry gas to demand nodes k (e.g., gas distribution companies) is bounded from below by constraint (28).

$$PCP_{p,t} \geq kc_p \sum_{k \in K} TP_{p,k,t} \qquad \forall p \in P, t > 1 \tag{28}$$

Moreover, compressor stations can be expanded in the planning horizon by installing new compressors at the same node. Eqs. (29) and (30) determine the total power of compressors installed up to time t at nodes j and p, respectively, and where τc is the compressor installation lead time in quarters.

$$JCP_{j,t} = JCP_{j,t-1} + JCInst_{j,t-\tau c} \qquad \forall j \in J, t > 1 \tag{29}$$

$$PCP_{p,t} = PCP_{p,t-1} + PCInst_{p,t-\tau c} \qquad \forall p \in P, t > 1 \tag{30}$$

3.1.4 Water Supplies

Water Demand for Drilling and Fracturing Wells. As explained before, a large amount of freshwater is required in the shale gas industry for the hydraulic fracturing of new wells. This model assumes that

the total amount of water required by a single well during the drilling, fracturing and completion processes (wr_i) is known (typically, 20 Mm^3/well) but may depend on the well location. Eq. (31) states that the total number of wells drilled, fractured and completed in pad *i* during period *t* determines the total water requirement of that pad at that period, and such amount should be supplied from one of more freshwater sources *f*. The amount of freshwater supplied by source *f* for drilling and fracturing new wells in pad *i* during period *t* is a key model decision represented by the continuous variable $WS_{f,i,t}$. In addition, if the well-pad *i* has the infrastructure for flowback water treatment and reuse, a reuse factor rf_i (usually below 20%) can reduce the need for freshwater, as shown in the LHS of eq. (31).

$$N_{i,t}\ wr_i\ /\ (1+rf_i)\ \ = \sum_{f\in F} WS_{f,i,t} \quad \forall i \in I, t \in T \tag{31}$$

Water Availability. Every freshwater resource (rivers, lakes, underground water, etc.) usually has an upper limit on the amount of water that it can provide to the shale gas industry, often given by a seasonal profile. If the parameter $fwa_{f,t}$ stands for the maximum volume of freshwater that source *f* can supply to the drilling and fracturing of new wells during the whole period *t*, the total amount supplied from *f* to every pad *i* should be bounded by above as in constraint (32).

$$\sum_{i\in I} WS_{f,i,t} \le fwa_{f,t} \qquad \forall f \in F, t \in T \tag{32}$$

3.1.5 Maximum Demands

A critical model decision is where to sell both the dry gas and the ethane flows produced by the shale gas processing plants. Every potential market (or demand node) is assumed to consume a maximum amount of product (dry gas for gas distributors, ethane for petrochemical plants) based on their own transportation or processing capacities. Moreover, such demand profile can be seasonal, especially in dry gas markets. Constraints (33) and (34) restrict the total flow of dry gas and ethane that can be sent from processing plants to each demand node during every period of the planning horizon.

$$\sum_{p\in P} TP_{p,k,t} \le gasdem_{k,t} \qquad \forall k \in K, t > 1 \tag{33}$$

$$\sum_{p\in P} LP_{p,l,t} \le ethdem_{l,t} \qquad \forall l \in L, t > 1 \tag{34}$$

3.2 Objective Function

The objective function of the model is to maximize the Net Present Value (NPV) of the long-term planning project as expressed in eq. (35).

$$\begin{aligned} NPV = \sum_{t\in T}(1+dr/4)^{-t} \Big[& \sum_{p\in P}\sum_{k\in K} gasp_{k,t}\ nd_t\ TP_{p,k,t} + \sum_{p\in P}\sum_{l\in L} ethp_{l,t}\ nd_t\ LP_{p,l,t} + \sum_{p\in P} lpgp_t\ nd_t\ NP_{p,t} \\ & - \sum_{i\in I}\sum_{j\in J} shgc_{i,t}\ nd_t\ FP_{i,j,t} \\ & - \sum_{p\in P} ks\ SepInst_{p,t}^{SepExp} \\ & - \sum_{i\in I}\sum_{n=1}^{\overline{n}_i} kd\ n^{WellExp}\ y_{i,n,t} \\ & - \sum_{i\in I}\sum_{j\in J} kp\ l_{i,j}\ DFP_{i,j,t}^{GasPipeExp} - \sum_{j\in J}\sum_{p\in P} kp\ l_{j,p}\ DGP_{j,p,t}^{GasPipeExp} \\ & - \sum_{p\in P}\sum_{k\in K} kp\ l_{p,k}\ DTP_{p,k,t}^{GasPipeExp} - \sum_{p\in P}\sum_{l\in L} kp\ l_{p,l}\ DLP_{p,l,t}^{LiqPipeExp} \\ & - \sum_{j\in J} kc\ JCInst_{j,t}^{CompExp} - \sum_{p\in P} kc\ PCInst_{p,t}^{CompExp} \\ & - \sum_{f\in F}\sum_{i\in I} \left(fix_f + var_f\ l_{f,i}\right) WS_{f,i,t} \Big] \end{aligned} \tag{35}$$

The objective function comprises positive and negative terms for every period of the planning horizon, discounted back to its present value by the annual discount rate of the project, *dr*. Positive terms are dry gas sales income, ethane sales income, and NGL other than ethane sales income. Negative terms are shale gas acquisition cost (including production, transportation and other operating costs), the cost of drilling, hydraulic fracturing and completion of shale gas wells, the cost of installing/expanding shale gas processing capacity at separation plants, the cost of constructing new pipelines either for

gathering raw gas or distributing dry gas and ethane, the cost of installing new compressor stations at junction nodes and processing plants, and freshwater acquisition and transportation costs for drilling and fracturing purposes.

It should be noticed that instead of using linear costs with fixed charges, nonlinear expressions are used to represent economies of scale functions in some of the negative terms of eq. (35), featuring exponents between 0 and 1. Hence, the objective function can be classified as non-convex, with strictly concave separable terms. However, all constraints are linear as was shown in the previous sections.

3.3 Cost Estimation

Special attention must be paid to the equipment costing in the objective function (35). Regarding shale gas separation plants and gas compressors, typical values for the exponents *SepExp* and *CompExp* vary from 0.60 to 0.77.[24] However, a particular case arises in this model for pipeline construction. By assumption 13, the cost of pipelines also follows an economy of scale function with regards to the pipeline diameter, with a typical exponent of 0.60. However, it should be noticed that pipeline diameters are not directly considered in the model but through the substituted variables $DFP_{i,j,t}$, $DGP_{j,p,}$, $DTP_{p,k,t}$ (for gas pipelines) and $DLP_{p,l,t}$ (for liquid pipelines). In fact, such variables account for the diameters raised to the power of 2.667 in the case of gas pipelines, and power 2 in the case of liquid pipelines. Therefore, if 0.60 is considered as exponent for the economy of scale regarding pipeline construction, the values of the exponents *GasPipeExp* and *LiqPipeExp* in the objective function (35) will be 0.60/2.667 = 0.225 and 0.60/2 = 0.30, respectively (see Appendix).

3.4 Model Adaptation to Account for Shale Gas Composition Variations

Relaxing the assumption of uniform shale gas composition so that the gas wetness is dependent of the well location significantly complicates the nature of the constraints. In order to precisely trace the composition of shale gas flows, the critical points in the proposed network superstructure are the junction nodes. If it is required to split flows to more than one separation plant, the model involves

bilinear equations so that the composition of all the outgoing flows takes a common value given that junctions are mixing-splitting nodes as stated by eqs. (36), (37) and (38).

$$GComp_{j,t}\ GP_{j,p,t} = GP^{G}{}_{j,p,t} \qquad \forall j \in J, p \in P, t > 1 \tag{36}$$

$$EComp_{j,t}\ GP_{j,p,t} = GP^{E}{}_{j,p,t} \qquad \forall j \in J, p \in P, t > 1 \tag{37}$$

$$LComp_{j,t}\ GP_{j,p,t} = GP^{L}{}_{j,p,t} \qquad \forall j \in J, p \in P, t > 1 \tag{38}$$

Eqs. (36), (37) and (38) include the additional variables $GComp_{j,t}$, $EComp_{j,t}$, $LComp_{j,t}$ (not dependent on the index p) which are the volume compositions of methane, ethane and LPG (hydrocarbons other than methane and ethane) in the shale gas, forcing all the flows departing from the junction node j (a mixing-splitting node) to have the same composition. Moreover, individual component flow balances are incorporated to the formulation through eqs. (39), (40) and (41).

$$\sum_{i \in I} FP^{G}{}_{i,j,t} = \sum_{p \in P} GP^{G}{}_{j,p,t} \qquad \forall j \in J, t > 1 \tag{39}$$

$$\sum_{i \in I} FP^{E}{}_{i,j,t} = \sum_{p \in P} GP^{E}{}_{j,p,t} \qquad \forall j \in J, t > 1 \tag{40}$$

$$\sum_{i \in I} FP^{L}{}_{i,j,t} = \sum_{p \in P} GP^{L}{}_{j,p,t} \qquad \forall j \in J, t > 1 \tag{41}$$

It can be easily seen that equations (36), (37) and (38) involve bilinear terms that add significant difficulty to the MINLP model, especially because the feasible region can no longer be modeled with linear constraints. However, the next section presents a particular case in which no bilinear terms have to be added when incorporating shale gas composition variations.

3.4.1 Shale Gas Flows Converging to a Single Processing Plant

If the model is intended to select only one of the given locations to install a separation plant (as it is expected due to the high cost of this kind of plants), linear expressions hold since no splitting occurs at

junction nodes. The tendency of the model to select only one plant location is demonstrated in the results section with Example 1.

Under this assumption the model modifications necessary to comply with shale gas composition variations are as follows. First, we include a new binary variable w_p, representing whether the location p is selected to install the plant. As a result, the single plant condition leads to constraints (42) and (43).

$$SepCap_{p,t} \leq sepmax\ w_p \qquad \forall p \in P, t > 1 \tag{42}$$

$$\sum_{p \in P} w_p \leq 1 \tag{43}$$

where *sepmax* is an upper bound on the capacity of a single gas processing plant. Note that although the plant location must be unique, it may be installed and expanded in different time periods.

In this way, upper bounds on the individual product flows emerging from every plant are imposed by constraints (44), (45) and (46) in place of eqs. (10), (11) and (12).

$$\max_{i \in I}\{gc_i\} \sum_{j \in J} GP_{j,p,t} \geq \sum_{k \in K} TP_{p,k,t} \qquad \forall p \in P, t > 1 \tag{44}$$

$$s_g^E \max_{i \in I}\{ec_i\} \sum_{j \in J} GP_{j,p,t} \geq \sum_{l \in L} LP_{p,l,t} \qquad \forall p \in P, t > 1 \tag{45}$$

$$s_g^L \max_{i \in I}\{lc_i\} \sum_{j \in J} GP_{j,p,t} \geq NP_{p,t} \qquad \forall p \in P, t > 1 \tag{46}$$

gc_i, ec_i and lc_i are the volume compositions of methane, ethane and LPG in the shale gas produced at pad i. Note that from eq. (14), if the plant is not selected (zero capacity) no shale gas flows can be sent to it, and the LHS of the last inequalities is zero. Finally, individual component balances are given by eqs. (47), (48) and (49). Eq. (47) for the methane balance is illustrated through the simple example depicted in Figure 4, comprising two well-pads, one junction node, the processing plant and the gas demand node.

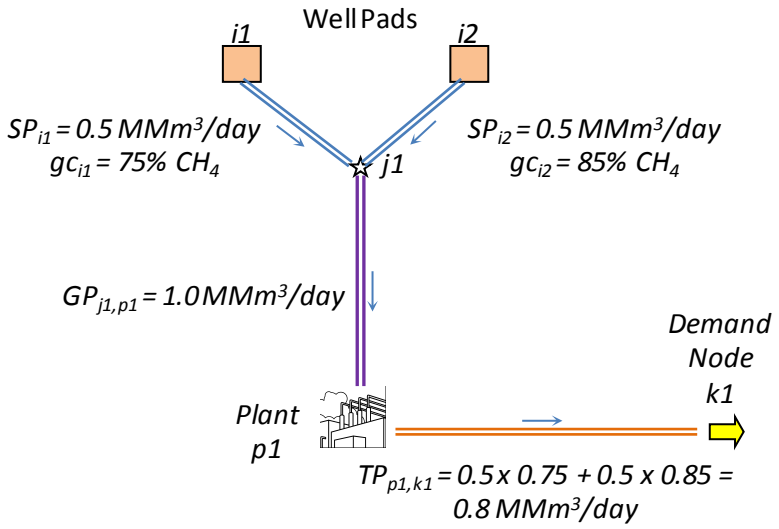

Figure 4. Mixing flows at a single processing plant.

$$\sum_{i \in I} gc_i \; SP_{i,t} = \sum_{p \in P} \sum_{k \in K} TP_{p,k,t} \tag{47}$$

$$s_g^E \sum_{i \in I} ec_i \; SP_{i,t} = \sum_{p \in P} \sum_{l \in L} LP_{p,l,t} \tag{48}$$

$$s_g^L \sum_{i \in I} lc_i \; SP_{i,t} = \sum_{p \in P} NP_{p,t} \tag{49}$$

In summary, the modified MINLP model accounting for shale gas composition variations according to the well site, assuming that a unique processing plant location is to be selected, seeks to minimize equation (35) subject to constraints (1)-(4), (8), (9), (13)-(34), (42)-(49).

4. SOLUTION STRATEGIES

Solving the MINLP models described in the previous sections is a very challenging task due to three main reasons: (1) the size of the model is large, (2) the objective function involves non-concave terms accounting for equipment costs (processing plants, pipelines and compressors), and (3) such non-linear functions have unbounded derivatives at zero values. The last two features directly follow from the economies of scale, usually used to model the equipment cost variation with regards to the equipment

size. In principle the MINLP model can be solved to global optimality with a spatial branch and bound search method with the use of convex envelopes for the concave terms in the objective (the secant). However, given the large size of the MINLP, the problem is intractable with such methods like the ones implemented in BARON, LINDOGLOBAL and COUENNE.[25] Therefore, in this section a tailored strategy is described for solving the large-scale MINLP problem.

4.1 Plant and Equipment Cost Estimations

Nonconvex power law expressions of the form $f(x) = c\, x^r$ with exponents less than one as in Biegler et al.[23] are commonly handled with two approaches: (a) approximate the concave function by a piece-wise linear function,[26] and (b) adding a small value ε to the variable x thus slightly displacing the curve so as to ensure non-zero argument in the function. Approximation (a) is computationally costly, but can be useful for generating global upper bounds for the maximization problem by solving approximate MILP problems with piecewise linear approximations (underestimations) of the concave equipment cost functions.[27][28] On the other hand, approximation (b) is meant to avoid unbounded derivatives but can have drawbacks, especially if the exponents are rather small.[29] To overcome this problem, a simple expression of logarithmic form is used here. In the following sections, both the piecewise linear approach and the logarithmic approximation used are briefly presented.

4.2 Piecewise Linear Approximation of Concave Cost Functions

Given the nonlinear concave cost functions in eq. (1) (generically referred to as $f(x)$), accounting for the cost $f(x)$ of a processing plant, pipeline, or compressor of size $x \in X$, it is simple to demonstrate that piecewise linear approximations like the one depicted in Figure 5 provide valid underestimations of $f(x)$. That is achieved by partitioning the domain of variable x into intervals ($X = [a_0; a_1] \cup [a_1; a_2] \cup \ldots [a_{m-1}; a_m]$ and introducing binary variables z_v to determine to what interval the selected value of x belongs. At every interval v, function $f(x)$ is approximated by: $\phi(x) = f(x_{v-1}) + (x - x_{v-1})\,(f(x_v) - f(x_{v-1})) / (x_v - x_{v-1})$. According to Padberg,[30] such a piecewise linearization can be modeled through two formulations: δ and λ. In this case, we adopt the δ-formulation that leads to:

$$
\begin{aligned}
& x = a_0 + \sum_{v=1}^{m} y_v \\
& \phi(x) = f(a_0) + \sum_{v=1}^{m} y_v \, [f(a_v) - f(a_{v-1})] / [a_v - a_{v-1}] \\
& y_v \geq (a_v - a_{v-1}) z_{v+1} \qquad v = 1 \ldots m-1 \\
& y_v \leq (a_v - a_{v-1}) z_v \qquad v = 2 \ldots m \\
& y_1 \leq a_1 - a_0 \\
& y_v \geq 0 \qquad v = 1 \ldots m \\
& z_v \in \{0, 1\} \qquad v = 1 \ldots m
\end{aligned}
\tag{50}
$$

Note that if $z_v = 0$ (with $v > 1$), $y_v = 0$ because of the fourth constraint in (50). From that, z_{v+1} is also zero to satisfy the third constraint in (50), given that $(a_v - a_{v\text{-}1})$ is always greater than zero. In other words, $z_v = 0$ implies $z_{v+1} = 0$, and equivalently, $z_{v+1} = 1$ implies $z_v = 1$, for $v > 1$. To illustrate the meaning of the variables in (50) reconsider the example given in Figure 5. Assume that $x = 6.5$. From the constraints described above, it follows that $z_2 = z_3 = z_4 = 1$, $y_1 = a_1 - a_0 = 2$ (because $z_2 = 1$), $y_2 = a_2 - a_1 = 2$ (because $z_3 = 1$), $y_3 = a_3 - a_2 = 2$ (because $z_4 = 1$), and $y_4 = 0.5$.

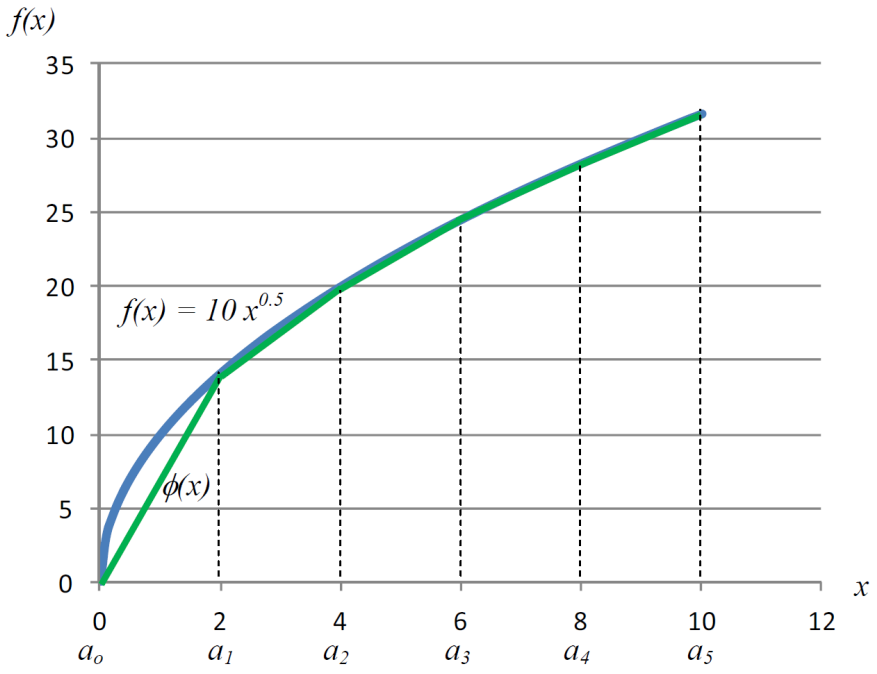

Figure 5. Concave cost function and piecewise linear underestimation.

By replacing the nonlinear terms in the objective function (35) with the piecewise linear approximations given in (50), the MINLP model reduces into an MILP model yielding valid upper bounds for the global optimum of the original problem. A key decision is how to divide the variable domain, i.e. how many intervals to consider. The finer the domain discretization, the closer is the upper bound to the actual objective value of the MINLP, but also the higher is the CPU time required by the MILP since the number of integer variables increases significantly. In this case, such a tradeoff is managed through the successive refining strategy presented in Section 4.4 based on the ideas of You and Grossmann[31][32] dealing with the nonlinear concave function $\sqrt{x}$.

4.3 Logarithmic Approximation of Concave Cost Functions

In order to avoid unbounded derivatives and estimation errors when solving NLP subproblems in the MINLP model, the alternate approximation function *g(x)* for *f(x)* by Cafaro and Grossmann[29] is used:

$$f(x) = c\, x^r \approx g(x) = k \ln(bx + 1) \tag{51}$$

where x is the size of the equipment, *f(x)* is the actual cost of the equipment of size x, *g(x)* is the estimated cost of the equipment of size x, and k , $b > 0$ are parameters selected to fit *f(x)* as closely as possible. Further details of this approximation are given in Cafaro and Grossmann.[29]

The proposed function has two main advantages with regards to the classic ε-approximation $f(x) \approx h(x) = c\,(x + \varepsilon)^r$: (1) the cost of $x = 0$ is exactly zero: $g(0) = k \ln(b\ 0 + 1) = k \ln(1) = 0$, and (2) the derivatives of *g(x)* for all $x \geq 0$ are bounded positive values given by $g'(x) = b\ k\ /\ (b\ x + 1)$. In particular at the origin ($x = 0$), $g'(x) = b\ k$. These properties are particularly useful when dealing with concave cost functions with small exponents, like the cost of pipelines (exponents 0.225 and 0.300, for gas and liquid pipelines, respectively).

Appropriate values for parameters k and b in function *g(x)* can be found relatively easily for liquid and gas pipelines, and the logarithmic approximation leads to very good results (less than 0.50 % error) in the calculation of pipeline costs in all of the case studies tackled in Section 5.

4.4 Solution Algorithm: Branch-Refine-Optimize (BRO) Strategy

To find the global optimum of the nonconvex MINLP model presented in Section 3, a two-level branch-and-refine procedure is proposed (see Figure 6). In the upper level we successively solve MILP approximations of the original MINLP problem following two purposes: (1) provide valid (and increasingly tighter) upper bounds of the global optimum, and (2) propose efficient supply chain network configurations. Once the MILP approximation is solved, the corresponding supply chain network design is fixed by removing all the nodes and arcs of the original superstructure not active in the MILP solution so as to define the lower level optimizing procedure.

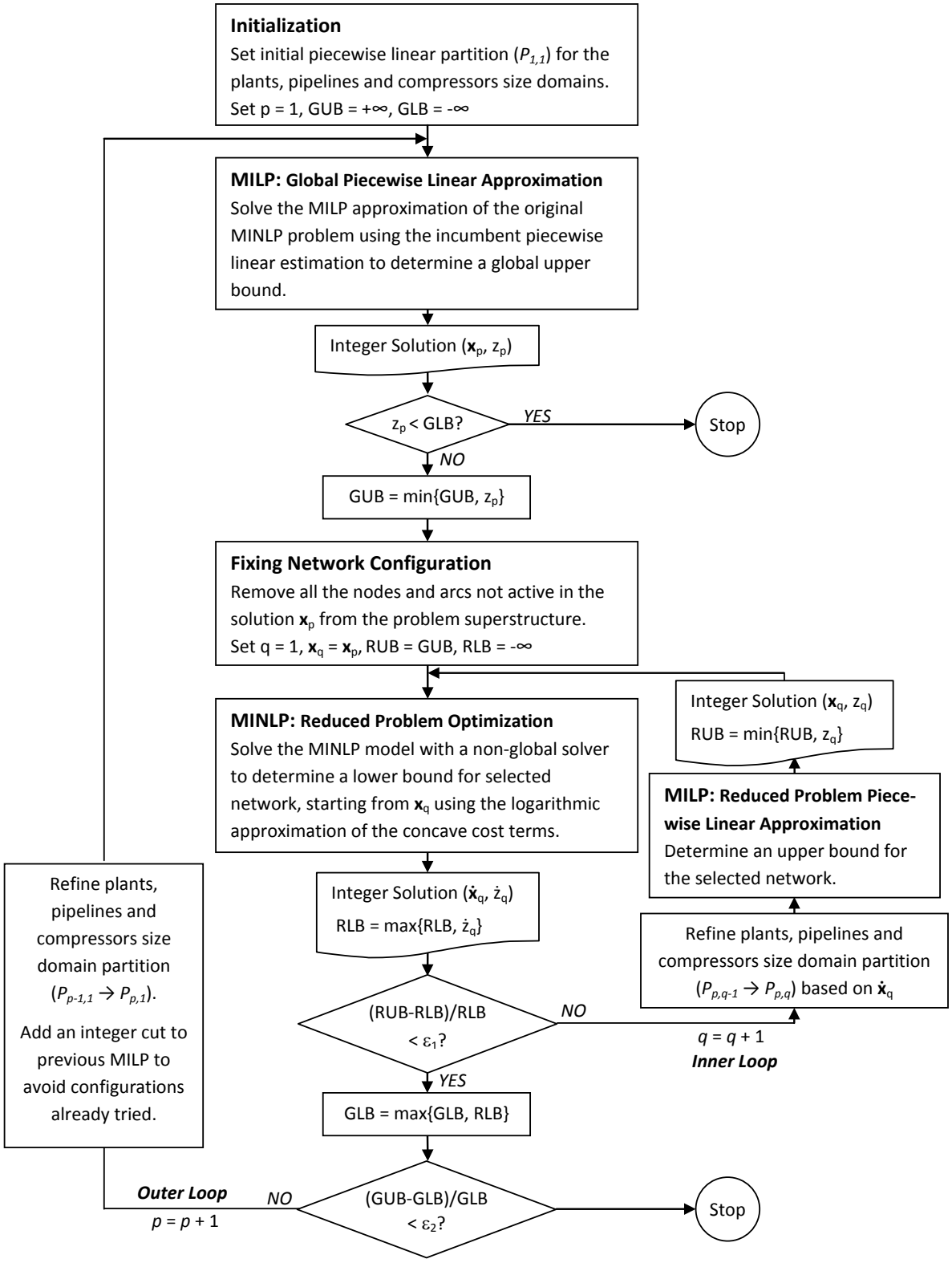

Figure 6. Branch-Refine-Optimize (BRO) algorithm.

The aim of the lower level of the algorithm is to find the global optimal solution of a reduced MINLP problem (or subproblem) focused only on equipment sizing (plant, pipelines and compressors) and the

drilling strategy (integer variables $n_{i,t}$) as the network structure is fixed. Since the reduced MINLP problem is nonconvex, its global optimal solution is found by solving on the one hand the reduced MINLP with a non-global solver (DICOPT, SBB)[25] to determine a lower bound for the selected network, and then successively partitioning the equipment size domains and recursively solving piecewise linear approximations of the objective function to determine tighter upper bounds (inner loop). Finally, the global optimal solution of the reduced MINLP is a feasible solution of the original MINLP, and its objective value provides a valid global lower bound of the problem.

Note that supply chain network designs proposed by the upper level at previous iterations are excluded with integer cuts in order to reduce the enumeration effort. Such integer cuts are similar to those proposed by Durán and Grossmann,[8] which eliminate particular binary combinations accounting for network configurations already analyzed. The cuts are derived from the values of the binary variables z_v used by the piecewise linear approximation of the concave cost terms in the objective function of the MILP (see Section 4.2). As a result, if the approximate solution obtained by the upper level in a new iteration is worse than the best solution found (or global lower bound), the algorithm automatically stops. Otherwise, the outer loop refines the piecewise linear approximation of the original problem and might improve the network structure so that the global optimal solution can be obtained after a finite number of iterations.

In summary, the proposed solution algorithm is as follows:

Step 1: Initialization. A one-piece linear underestimation (secant) is used for all the concave cost terms of later periods (for instance, t > 10), while in earlier periods the starting piecewise linearization comprises two to four intervals. The global upper bound is set to GUB = $+\infty$, and the global lower bound GLB = $-\infty$.

Step 2. Global Piecewise Linear Approximation. Solving the incumbent MILP approximation of the original MINLP problem (as shown in Section 4.2) provides a global upper bound GUB. Since all the constraints in the MINLP are linear, the optimal solution of the MILP is also a feasible solution of the MINLP problem. Thus, a global lower bound (GLB) can be directly obtained by substituting the

optimal solution of the MILP into the MINLP. However, this solution can be taken as the initial point of a non-global MINLP solving step to improve the GLB.

Step 3. Reduced Problem Optimization. By fixing the network structure, i.e. removing all the nodes (well-pads, junctions, gas processing plants, compressors) and arcs (pipelines) that were not selected in the optimal solution of the MILP, we successively solve a reduced MINLP problem with non-global algorithm (DICOPT, SBB) that is intended to improve the best solution found. The MINLP model makes use of the logarithmic approximation presented in Section 4.3, which avoids the numerical difficulties reported by You and Grossmann.[31] In this way, solving the nonconvex reduced MINLP might yield an improved lower bound for the subproblem (RLB). Next, based on the optimal values of the equipment size variables, we bisect the corresponding intervals of the piecewise linear approximations. If the optimal solution of the MINLP problem lies at the bounds of some intervals, we do not add a new interval for these terms. After refining the domain partition, we can obtain a tighter upper bound for the reduced problem (RUB), as shown in the next step.

Step 4. Reduced Problem Piecewise Linear Approximation. The MILP with the piecewise linear approximation of the reduced problem provides an upper bound RUB, whose value tends to decrease as the domain partition is refined. The inner optimization loop iterates until the lower bound from the MINLP and upper bound of the MILP are within an optimality tolerance ε_1. Once that occurs, the global optimum of the reduced problem has been found, and the lower bound of the original problem (GLB) is updated.

Step 5. Stopping Criteria. From the values of the variables in the best solution found by the reduced MINLP, the intervals of the piecewise linear approximations in the original problem are bisected, and a new integer cut is added to the upper level MILP to avoid network configurations already tried. Next, the algorithm returns to Step 2 and two cases may occur: (a) a tighter global upper bound is found, or (b) the approximate solution is worse than the global lower bound. In case (a), the main optimization loop keeps iterating until the global lower and upper bounds are close enough to satisfy the optimality criteria ε_2 ($> \varepsilon_1$). In case (b), the algorithm stops and the optimal solution is the best solution found.

5. RESULTS AND DISCUSSION

In order to illustrate the application of the MINLP model and the proposed optimization algorithm, three examples are considered in this section. Example 1 deals with a real-size illustrative problem for optimizing the supply chain network design for a new shale gas exploitation area covering more than 150,000 km^2. In this case, a different production profile is assumed for each potential site where the wells are drilled. However, the gas "wetness" (or hydrocarbon composition) is assumed to be the same in each well pad. In turn, Example 2 is a variant of the previous case where the gas wetness becomes dependent on the well location. The aim of the second example is twofold: (a) find out how the gas wetness distribution affects the drilling strategy, and (b) highlight the contribution of the hydrocarbons other than methane to the economics of the project. The third example introduces variations in the pipeline pressures in order to show changes in the optimal solution. Finally, a real-world case study of the U.S. shale gas industry is tackled at the end of this section.

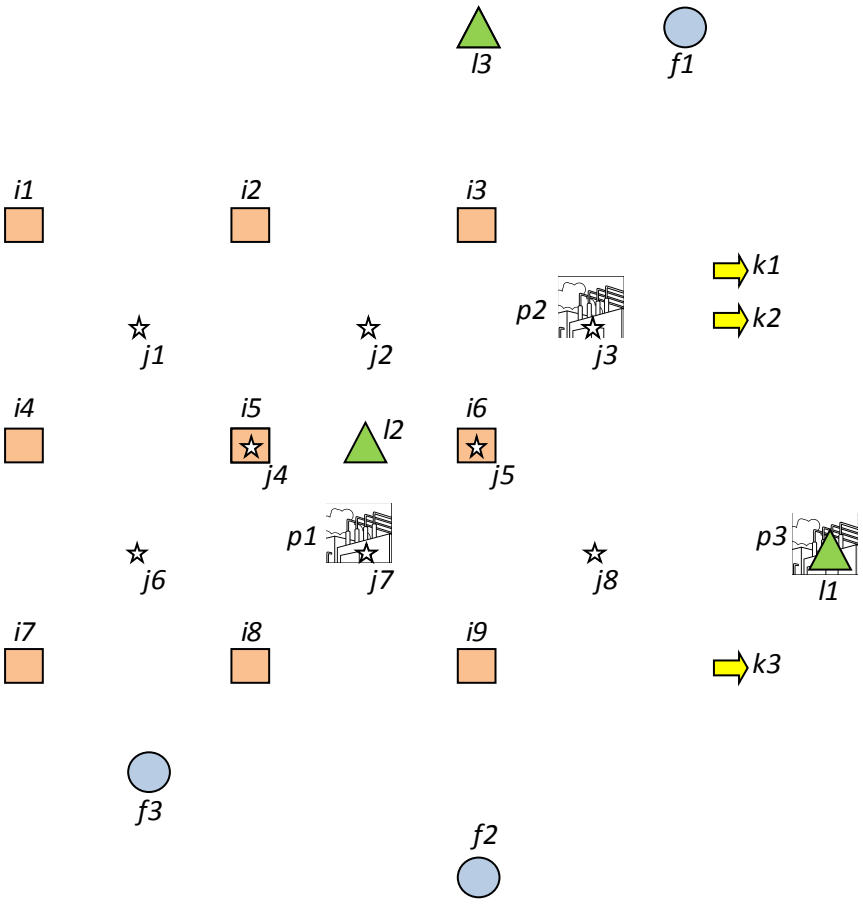

Figure 7. Nodes of the supply chain network superstructure for Examples 1, 2 and 3.

5.1 Example 1: The same shale gas wetness in all the wells.

Consider the shale gas supply chain superstructure whose nodes are shown in Figure 7. It comprises nine potential sites for drilling wells (*i1*...*i9*), eight potential sites for junction/compression nodes (*j1*..*j8*), three possible sites for processing plant installation (*p1*...*p3*), three methane demanding nodes (*k1*...*k3*), three ethane demanding nodes (*l1*...*l3*), and three freshwater sources (*f1*..*f3*). The Cartesian coordinates of each site (in km) are given in Table 1. Distances between nodes for pipeline length and water transportation calculations are measured in Euclidean norm.

Table 1. Cartesian Coordinates of Problem Nodes (in km)

	well pads									*junction nodes*							
	i1	*i2*	*i3*	*i4*	*i5*	*i6*	*i7*	*i8*	*i9*	*j1*	*j2*	*j3*	*j4*	*j5*	*j6*	*j7*	*j8*
x	0	50	100	0	50	100	0	50	100	25	75	125	50	100	25	75	125
y	0	0	0	50	50	50	100	100	100	25	25	25	50	50	75	75	75

	processing plants			*methane demand nodes*			*ethane demand nodes*			*freshwater sources*		
	p1	*p2*	*p3*	*k1*	*k2*	*k3*	*l1*	*l2*	*l3*	*f1*	*f2*	*f3*
x	75	125	175	145	145	145	175	75	100	150	100	25
y	75	25	75	15	25	100	75	50	-50	-50	150	125

The planning horizon comprises 40 time periods (quarters) and the annual rate that was considered for discounting back cash flows is 13.5%. The methane price is assumed to be seasonal, with a base price of \$142.86/Mm^3 for periods *t1*, *t5*, *t9*,…, and seasonality factors of 1.10 for periods *t2*, *t6*, *t10*,…; 1.25 for *t3*, *t7*, *t11*,…; and 1.10 for *t4*, *t8*, *t12*,… The shale gas cost is fixed at \$35.71/Mm^3, the price of liquid ethane is \$329.48/ton, and other hydrocarbons (heavier than ethane) are liquefied petroleum gases (propane, butanes and pentanes) separately sold at \$749.56/ton (more than double of the ethane price). The liquid ethane density is 0.546 ton/m^3, while the LPG density is averaged at 0.600 ton/m^3. Maximum methane demands are 10, 5 and 15 MMm^3/day for nodes *k1*, *k2* and *k3*, while maximum ethane demands

at nodes *l1*, *l2* and *l3* are 2500, 2000 and 1500 ton/day, respectively. LPG maximum demands are 3000 ton/day at every node *p*.

Freshwater availability is also assumed to be seasonal, with reference values of 250, 80 and 190 Mm^3/quarter for sources *f1*, *f2* and *f3*, respectively, and seasonality factors of 1.20 for periods *t1*, *t5*, *t9*,…; 1.00 for periods *t2*, *t6*, *t10*,…; 0.80 for *t3*, *t7*, *t11*,…; and 1.10 for *t4*, *t8*, *t12*,…. Every individual well requires 20 Mm^3 to be drilled and hydraulically fractured regardless of its location. Moreover, no more than three wells can be drilled in a single location during one quarter, and a total of 20 wells is the maximum number permitted for a single well-pad. Overall, a total of 180 wells can be drilled over the time horizon.

The shale gas pressure at well-pads is set to 2.1 MPa, compressors at junction nodes increase the shale gas pressure from 1.4 to 2.1 MPa, processing plants receive the shale gas at 1.4 MPa, while compressors increase the methane pressure from 4.0 to 6.0 MPa. Finally, methane is delivered at demand nodes at 4.0 MPa.

Regarding the cost of processing plants, wells and compressors, economies of scale functions of the form $C(x) = c\ x^r$ are used, with c = MM\$210, MM\$5, MM\$0.011150, and r = 0.60, 0.60, 0.77, respectively. The units of the size variables are MMm^3/day for plants, wells for drilling/fracturing, and kW for compressors. For costing pipelines, a function $C(l,D) = c\ l\ D^r$ is used, with c = MM\$0.125594, r = 0.60, l (length) measured in km and D (diameter) in inches. The same function is used regardless of the product transported (liquid or gas) and the nodes being joined. For instance, the cost of a pipeline of 10 inches in diameter and 100 km in length is MM\$50.

The shale gas productivity (in MMm^3/day) at every well is modeled as a decreasing function of the well age t (see Figure 2) with the form $P(t) = k_i\ t^{-0.37}$, for t = 1…40. The constant k_i is 0.0806 for wells drilled in locations i = *i1*, *i4*; 0.0732 for i = *i2*, *i5*, *i7*; 0.0659 for i = *i3*, *i6*, *i8*; and 0.0586 for i = *i9*. Finally, the shale gas composition (independent of the well location) and its water content are given in the second column of Table 2. Notice the relatively high composition of wet gas (about 25%, with half of it being ethane).

Table 2. Shale Gas Water Content (kg/MMm^3) and Composition (Molar % in Dry Basis)

	Example 1	*Example 2*								
		i1	*i2*	*i3*	*i4*	*i5*	*i6*	*i7*	*i8*	*i9*
H_2O (kg/MMm^3)	615	615	615	615	615	615	615	615	615	615
N_2 (Mole %)	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
CO_2 (Mole %)	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
CH_4 (Mole %)	74.6	87.6	83.6	80.6	82.6	80.6	77.6	78.6	75.6	74.6
C_2H_6 (Mole %)	12.8	5.8	7.8	8.8	9.8	9.8	11.8	10.8	12.8	12.8
C_3H_8 (Mole %)	7.6	3.6	4.6	5.6	4.6	5.6	5.6	6.6	6.6	7.6
i-C_4H_{10} (Mole %)	1.2	0.5	0.7	1.2	0.5	0.8	1.1	0.7	0.8	1.2
n-C_4H_{10} (Mole %)	0.8	0.2	0.3	0.8	0.2	0.2	0.9	0.3	0.7	0.8
i-C_5H_{12} (Mole %)	0.5	0.2	0.5	0.5	0.2	0.7	0.5	0.7	1.0	0.5
n-C_5H_{12} (Mole %)	0.5	0.1	0.5	0.5	0.1	0.3	0.5	0.3	0.5	0.5

After implementing the BRO solution algorithm for this example, the optimal design for the shale gas supply chain is depicted in Figure 8 and the strategic drilling plan yields a NPV of MM$1664.48. The optimal solution determines that only one shale gas processing plant is installed at site *p1*, with a maximum capacity of 6.594 MMm^3 of shale gas per day and a total cost of MM$651.2. Due to the economies of scale, the plant and all the pipelines are installed in the first period of the time horizon with no expansions planned over the first ten years. Regarding shale gas compression power at junction nodes, 1236 kW, 706 kW and 818 kW are installed at nodes *j4*, *j5* and *j6*, in that order.

The selected destinations for methane and ethane are nodes *k3* and *l2*, respectively. Methane is supplied by a gas pipeline of 74.33 km length, and 17 ½ inches in diameter (or the upper closest diameter for gas commercial pipelines), requiring a compressor of 2428 kW. In turn, ethane is transported through a liquid pipeline of 5 ¾ inches in diameter. The maximum flow for both pipelines is reached in quarter *t7* when the plant is operated at full capacity to produce methane at the rate of 4.915 MMm^3/day, and ethane at 1130 ton/day (see Figure 9). The production level at the plant keeps high for other 6 quarters, until the maximum number of wells at every pad (20) is reached.

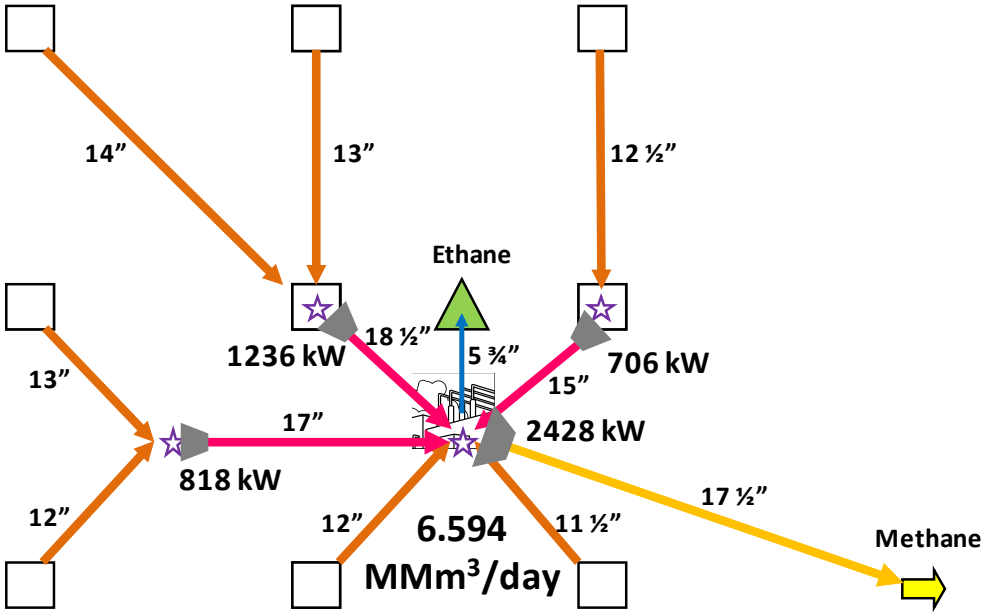

Figure 8. Optimal design for the shale gas supply chain network of Example 1.

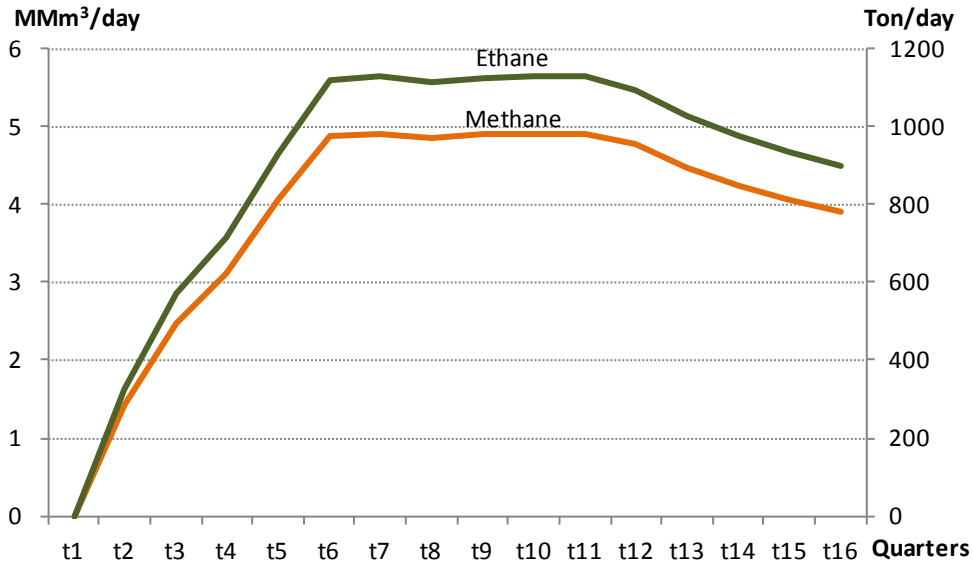

Figure 9. Amount of methane and ethane produced during the first four years in the optimal solution of Example 1.

One of the important features of the model is the ability to generate an optimal drilling strategy so as to keep the level of production well balanced over the entire time horizon (see Figure 9). In this way, plant, compressor and pipeline sizes can be smaller than those needed when a very intensive drilling plan is applied.

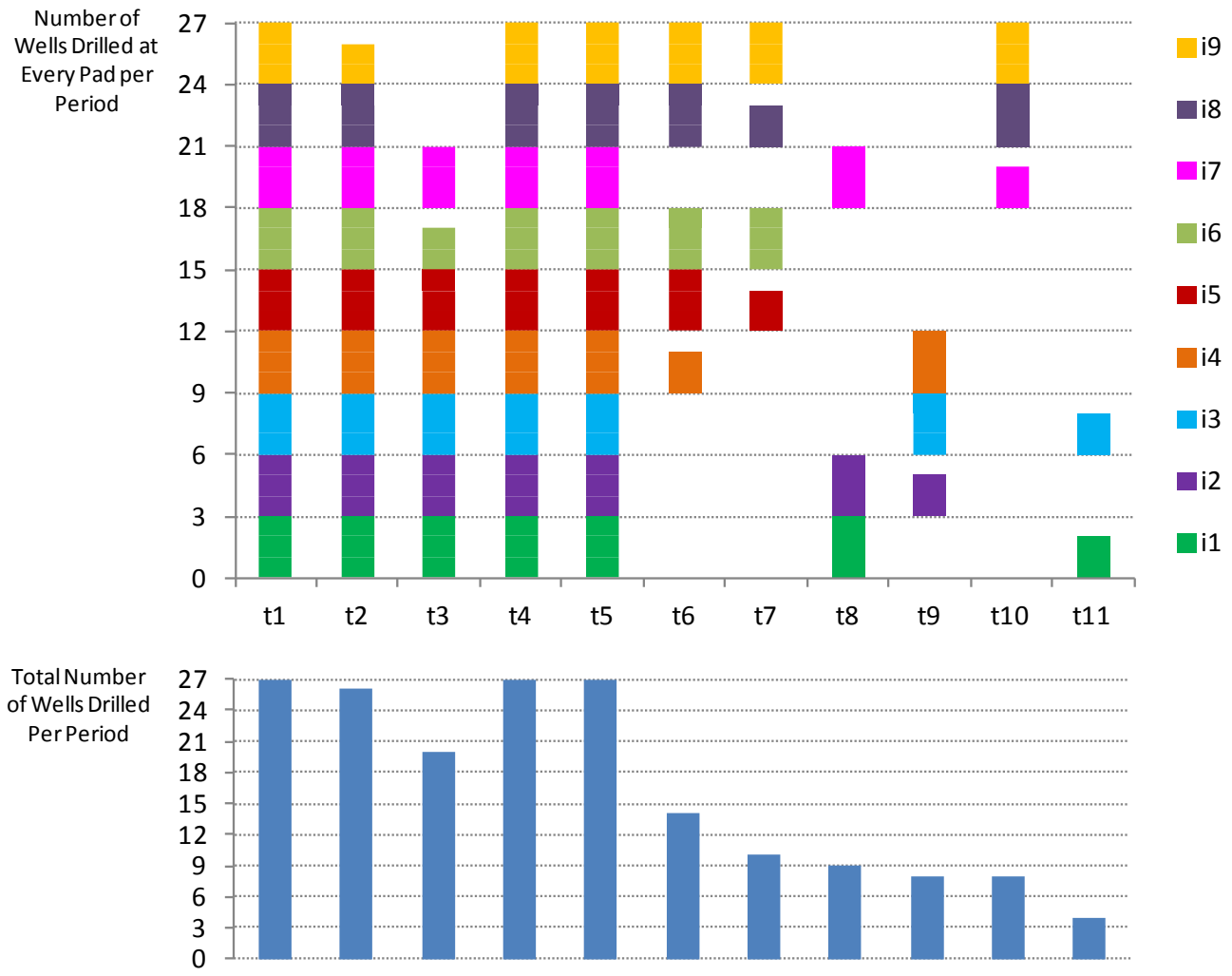

Figure 10. Optimal drilling strategy for Example 1.

The optimal drilling plan is depicted in Figure 10 in which each pad is represented with up to 3 wells that are drilled in a single time period. The height of each single-colored column at every line (which can be 0, 1, 2 or 3) represents the number of wells being drilled at each location in every period. The drilling plan is developed over the first three years of the planning horizon, and two phases can be easily distinguished: (1) Intensive drilling phase, and (2) Flow maintenance phase. The first phase covers the first five quarters, and its main objective is to drill and fracture as many wells as possible since there are no wells at the initial time. However, this strategy is partially limited by the water availability, which is scarce in periods *t2* and *t3*. Even under these circumstances, the model tends to rapidly increase the shale gas production focusing on the most productive regions. The second drilling phase takes place

during the following six quarters, and seeks to maintain a stable flow of shale gas in every pipeline until the maximum number of wells (20) is reached in every well-pad.

Overall, the optimal strategy yields a positive net present value of MM$1664.48, with a total investment in period *t1* amounting to MM$1077.60. Of the initial investment, 60% corresponds to the gas processing plant, 30% to pipeline installation, 8% to well drilling and fracturing costs, 1% to compressors, and less than 1% to water acquisition and transportation charges. Finally, the discounted payback period of the project is 3 years. Most of the project revenues come from LPG sales (50%) followed by methane (34%), and at last ethane (16%). The next example is proposed to analyze how the solution changes when the shale gas wetness depends on the well location, with the gas being much drier in some regions.

5.2 Example 2: Variable shale gas wetness.

The second example is a variant of Example 1 in which the shale gas wetness is dependent on the well site. The shale gas composition with regard to the location is presented in Table 2 where it can be seen that the composition of wet gas is less than 25% in many well pads. The only site producing shale gas with exactly the same composition as in Example 1 is node *i9*. The shale gas becomes drier in the direction of node *i1*. In fact, the methane mole percentage increases from 74.6% (node *i9*) to 87.6% (node *i1*). All the other data remain unchanged. Regarding the MINLP model, we use the modified version of the model presented in Section 3.4.1, which preserves linearity in the constraints under the assumption that a single processing plant is installed. Comparing the solution with the one obtained in Example 1, it can be concluded that the assumption of a single processing plant installation is not such a restrictive assumption for our case study.

In fact, the optimal network configuration obtained is exactly the same as for Example 1. This is an expected result, since the total amount of shale gas produced in every pad is the same. There are only minor variations in the pipeline diameters, compressor power and processing plant size. In particular, the plant capacity is reduced from 6.594 MMm^3/day to 6.222 MMm^3/day due to a more extended drilling

strategy. The main differences in the shale gas composition definitively affect the drilling strategy, as well as the economics of the project. As shown in Figure 11, the optimal drilling strategy now tends to prioritize the pads producing wetter gas (i.e., those producing a higher amount of heavier hydrocarbons). Wells drilled in less attractive pads (*i1, i2, i3*, *i4, i5*) are left for later periods (*t8* to *t13*). As a result, the overall drilling strategy now takes 13 periods instead of 11.

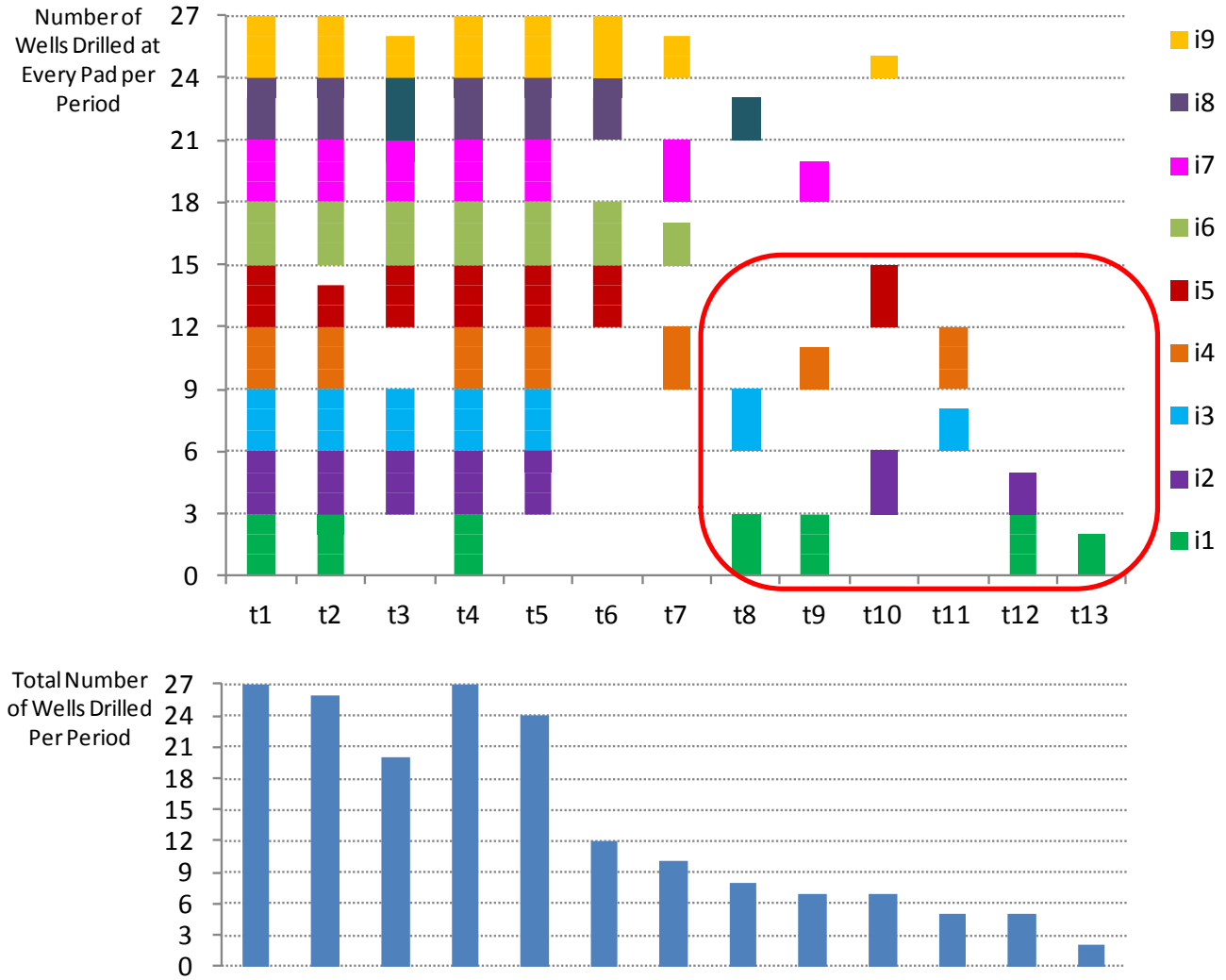

Figure 11. Optimal drilling strategy for Example 2.

From Figure 12 it can be seen that the production of methane is extended through a longer period of time, but the amount of ethane and heavier hydrocarbons is significantly lower than in Example 1. In summary, the optimal strategic plan involves an initial investment of MM$1054.67, while the net

present value is MM$1202.54, 27.8% below the net present value for Example 1. As a consequence, the discounted payback period increases from 3 to 3.7 years. The economic differences can be clearly noticed in the product sales income distribution. Given the new shale gas composition for each well-pad, LPG and methane sales represent 43% of the total income, while ethane is only 14%; versus 34% and 16% in the previous example.

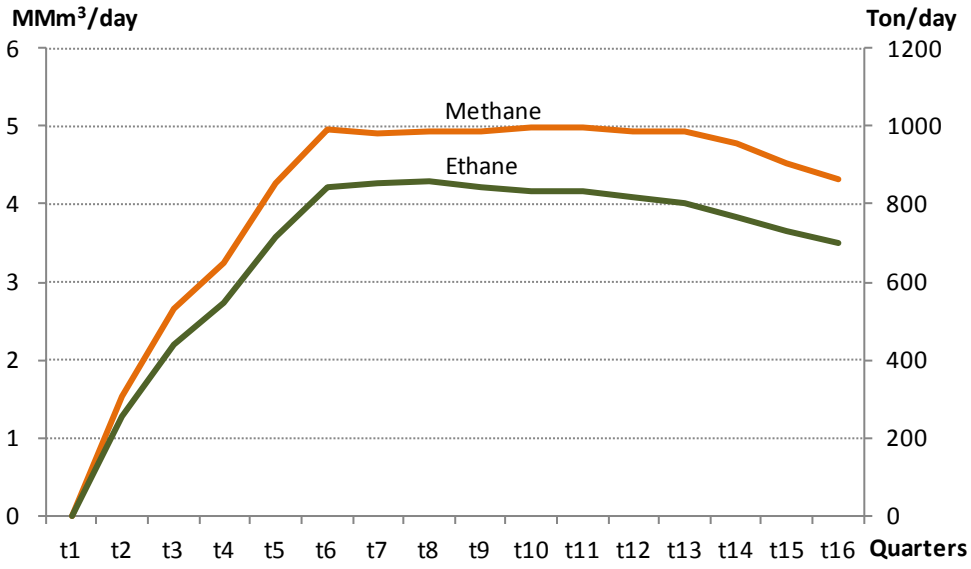

Figure 12. Amount of methane and ethane produced during the first four years in the optimal solution of Example 2.

5.3 Example 3: Changes in gas pipeline pressures.

By assumption (9), all the gas pressures are specified at some fixed values before solving the model. In Examples 1 and 2, the shale gas pipeline pressures vary from 2.1 MPa (inlet pressure) to 1.4 MPa (outlet pressure), while transmission pipelines transport dry gas from 6.0 MPa to 4.0 MPa. In both cases, gas compressors are assumed to operate at a common compression ratio of 1.5 (a typical value for centrifugal compressors). Even though determining the optimal pressure for every pipeline is out of the scope for this model, Example 3 is intended to show how the results are affected by changes in the pressure values. Example 3 is a variation of Example 1 in which both shale gas and dry gas compressors operate at a pressure ratio of 2, while shale gas at wellbores is delivered at a pressure of 2.8 MPa instead

of 2.1 MPa. More precisely, gathering pipelines transport shale gas from 2.8 MPa to 1.4 MPa, while dry gas is transported through transmission pipelines from 8.0 to 4.0 MPa.

The main findings of this example are related to pipeline and compressor sizing, since the pipeline network structure, the processing plant size and the drilling strategy do not change in the optimal solution. As expected, pipeline diameters can be reduced at the expense of using higher power in the compressors. On the one hand, the pipeline diameters are reduced 15.75% (from 12.38 to 10.43 inches on average), the gathering pipelines are reduced by 15.91% (from 16.59 to 13.95 inches) and the gas transmission pipeline by 14.71% (from 17.33 to 14.78 inches). On the other hand, shale gas compressors (a total of three, at junction nodes *j4*, *j5* and *j6*) increase their total power by a factor of 1.71 (from 2760.47 kW to 4723.31 kW), while the only dry gas compressor at the outlet of the processing plant has a total power of 4189.29 kW (3000 kW installed in period *t1* and the remaining 1189.29 kW in period *t4*), which implies a 72.6% increase in the methane compressor power compared to Example 1.

Overall, the pipeline installation cost is reduced from MM$321.35 to MM$291.10, the investment in compressor stations increases from MM$11.14 to MM$17.56, while the NPV of the project is improved by 1.15%. Although the difference is rather small, future work will focus on determining the optimal pressures for the gas pipelines.

5.4 Computational Results

The most time-consuming step in the BRO algorithm is the solution of the MILP approximation of the full-size problem, i.e. the global piecewise linear approximation. As proposed in Section 4.4, we initialize the algorithm with a one-piece linear underestimation for all the concave cost terms for periods $t > 10$, while the starting piecewise linearization involves two to four intervals for $t \leq 10$ (two for pipelines and compressors, four for processing plants). Even under those conditions, the size of the first MILP approximation of Example 1 is rather large: 51,880 equations, 47,643 continuous variables, and 3,490 binary variables (2,343 after pre-processing), as can be seen in Table 3. From the latter, 1,440 determine the number of wells to drill at every period, while the remaining are the binaries of the δ piecewise linearization. Even though the relaxation is somewhat tight (16.8% integrality gap), the first

MILP takes almost 5 hours of CPU time (using GAMS/GUROBI 5.5.0[25] on an Intel Core i7 CPU, 2.93 GHz, 12 GB RAM, with 6 parallel threads) to solve the problem with an optimality gap of 0.25%. Having solved this problem, the first global upper bound is found: MM$1709.04. Next, the solution found is used as the initial point of a non-global MINLP optimization algorithm (GAMS/DICOPT 24.1.3, with GUROBI 5.5.0 as the MILP solver, and CONOPT 3.15 as the NLP solver)[25] which after 4 major iterations and 190 CPUs finds the first optimized integer solution, yielding an NPV = MM$1655.83 (3.21% of global optimality gap).

Table 3. Computational Results for Example 1.

Outer Loop #	Inner Loop #	MILP				MINLP					Red. Opt. Gap % (ε_1)	Global Opt. Gap % (ε_2)
		Cont. Var.	Bin. Var.	Eq.	CPUs	Cont. Var.	Bin. Var.	Eq.	Major Iters.	CPUs		
1	1	47,643	3,490	51,880	17,530	31,633	1,440	28,900	4	190	3.21	3.21
	2*	21,344	3,737	24,578	14	5,087	1,440	5,532	3	20	2.12	2.94
	3*	21,591	3,984	25,072	35	5,087	1,440	5,532	3	13	1.34	2.94
	4*	21,817	4,210	25,524	44	5,087	1,440	5,532	4	21	0.98	2.88
2†	1	47,644	3,491	51,883	30,922	31,633	1,440	28,900	5	364	2.27	2.27
	2*	21,344	3,737	24,578	20	5,087	1,440	5,532	4	25	1.60	2.27
	3*	21,591	3,984	25,072	17	5,087	1,440	5,532	3	7	1.14	2.27
	4*	21,798	4,191	25,486	25	5,087	1,440	5,532	3	7	0.81	2.27

* *Network configuration is fixed*

† *Plant cost estimation is refined*

At the next step, the BRO algorithm fixes the network configuration, and the inner loop starts to optimize the reduced MINLP problem, by successively refining piecewise linear approximations of the concave cost terms in the objective function. As observed in the second line of Table 3, the first reduced MILP problem has one half of the binaries of the full-size MILP approximation, while the number of

continuous variables and equations are cut down by a factor of 7 and 5, respectively. In fact, the MILP approximations of the reduced problem have small sizes, never requiring more than 60s to find the optimal solution (0.25% optimality gap). From Table 3 it follows that the optimality gap of the reduced MINLP problem falls below $\varepsilon_1 = 0.10$ after 4 iterations. At that point, the algorithm adds an integer cut removing the network configuration already tried, and refines the full-size MILP piecewise linearization based on the values of the variables at the best solution found.

Even though the size of the MILP does not increase considerably in the second iteration of the outer loop, the time to find the optimal solution increases to 8.5 hours. Figure 13 shows the progress of the global upper bound, the best solution found, and the upper bound for the solution of the reduced problem over two iterations of the outer loop of the BRO algorithm. Overall, after solving eight MILP and eight MINLP models in 13.7 h of CPU time, the global optimality gap is reduced below $\varepsilon_2 = 2.5\%$.

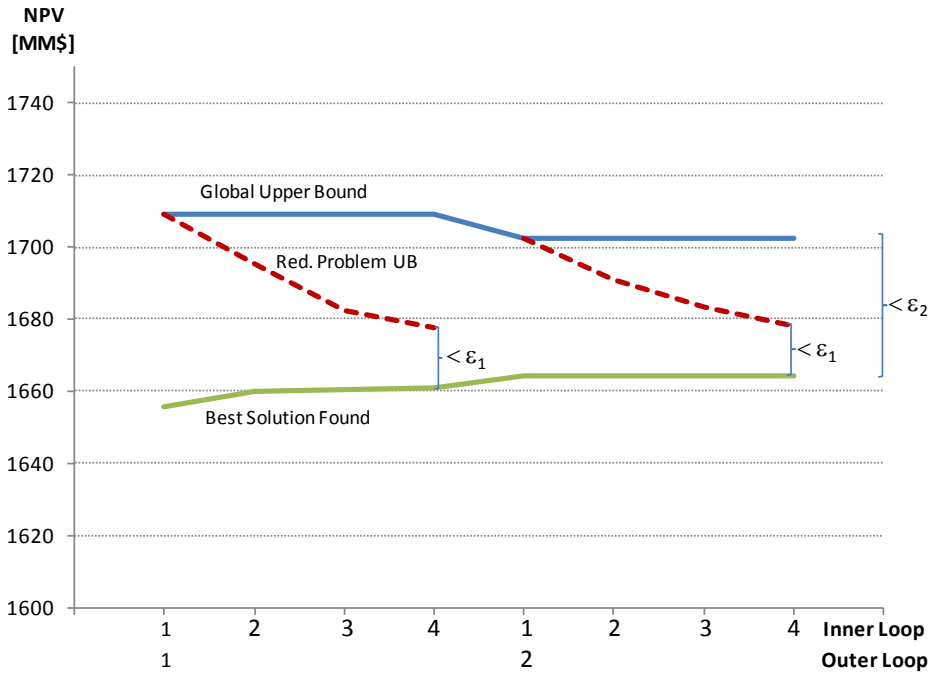

Figure 13. Progress of the global upper bound, the reduced problem upper bound and the best solution found in the solution of Example 1 through the BRO algorithm.

Regarding Example 2, the alternate formulation presented in Section 3.4.1 slightly increases the size of the models compared to Example 1. In the first iteration, the global MILP approximation has 51,998 constraints, 47,643 continuous variables, and 3,493 integer variables, taking more than 12h of computational time to reduce the optimality gap below 1.00%. After two iterations of the outer loop in more than 24h of computational time the global optimality gap is 7.5%.

5.5 Real-world case study: Shale Gas Development Project

A shale gas production company is interested in expanding the drilling and production activity in the Marcellus shale play. The company has determined more than 150 potential sites for well pads, which can be grouped into nine regions. All the shale gas produced in each region is collected by a low-pressure trunk pipeline that transports the gas to a nearby compressor station. Finally, the raw gas is dehydrated and sent through high pressure transmission pipelines tied to midstream lines owned by third party distribution companies. Pipeline construction and compressor installation require considerable lead-times (more than two years), which are considered in the formulation. In addition, the company has the possibility of drilling the wells and keeping them closed for some periods until the pipelines collecting the shale gas become available. Such an assumption requires a model modification shown in the Appendix. Besides, a maximum of four wells per pad can be drilled and completed in a single period (up to twelve wells in at most three pads per period), and each pad should not contain more than ten wells. Fourteen freshwater reservoirs are available in the area. Due to confidentiality reasons, further details on the problem cannot be given.

Since the shale gas is dry (95% mol methane) the study does not account for gas processing and fractionation plants. However, the large number of pads yields a large-scale MINLP model with 4,815 discrete variables, 12,226 continuous variables and 16,815 constraints. After two major iterations of the BRO algorithm and 71,000 CPUs of computational time, the optimal solution yields an NPV of MM$815, with a global optimality gap of 8.2%. The most convenient regions to be exploited during the following 10 years are regions 2 and 6, as seen in Figure 14, where a total of 22 and 18 pads are

constructed, respectively. Gathering pipelines of 7 to 10 inches in diameter collect the shale gas in each region, while trunk pipelines of 23 to 24 inches, and transmission pipelines of 12 inches (to delivery point 1) and 18 inches (to delivery point 2) are planned. Finally, a compressor station with a total power of 32,000 kW should be installed. Freshwater for drilling and fracturing is supplied by three of the available reservoirs. Figure 15 shows the drilling strategy for the 380 selected wells of regions 2 and 6, while Figure 16 illustrates the shale gas flows in major pipelines for periods 14 to 40, showing the trend of the model toward maximizing the pipeline utilization by maintaining a stable flow over time.

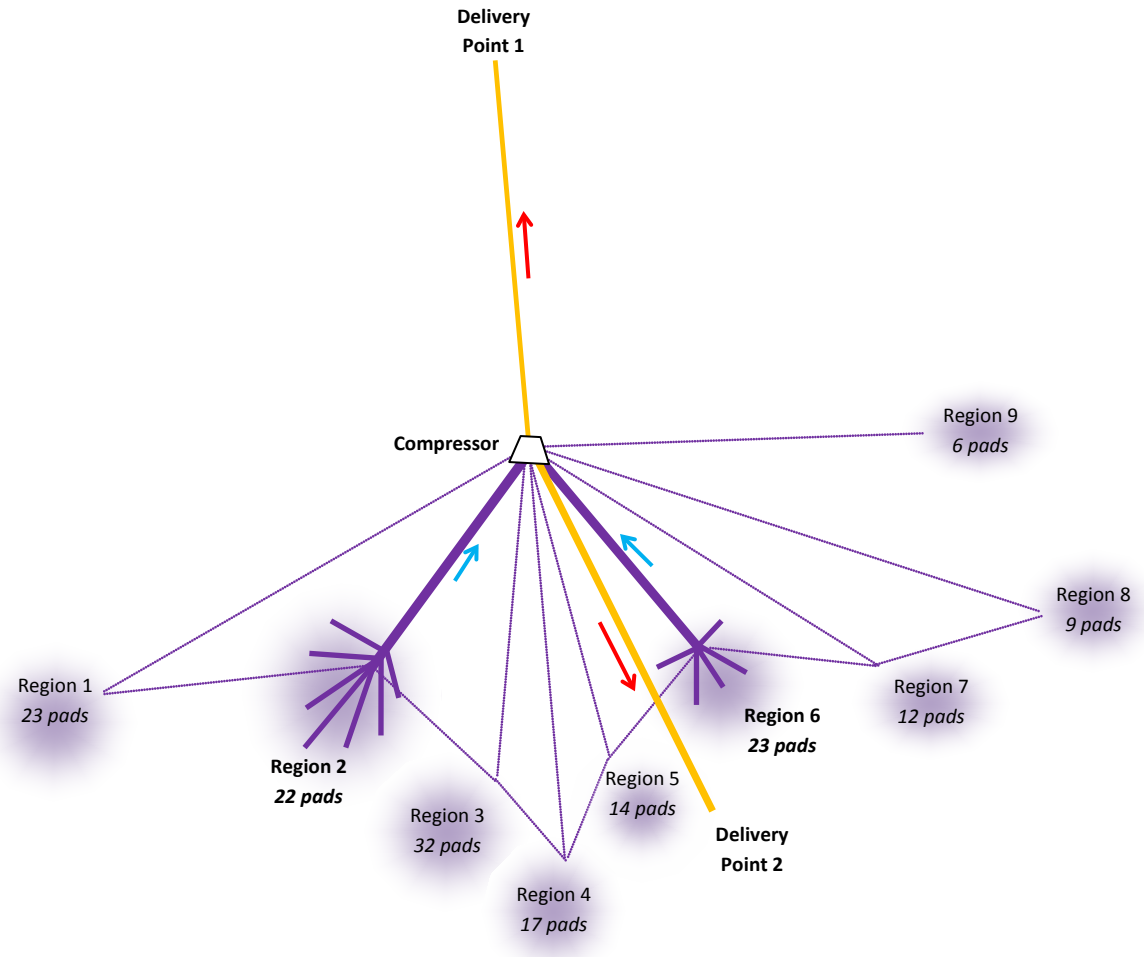

Figure 14. Schematic representation of the shale gas supply chain superstructure for the real-world case.

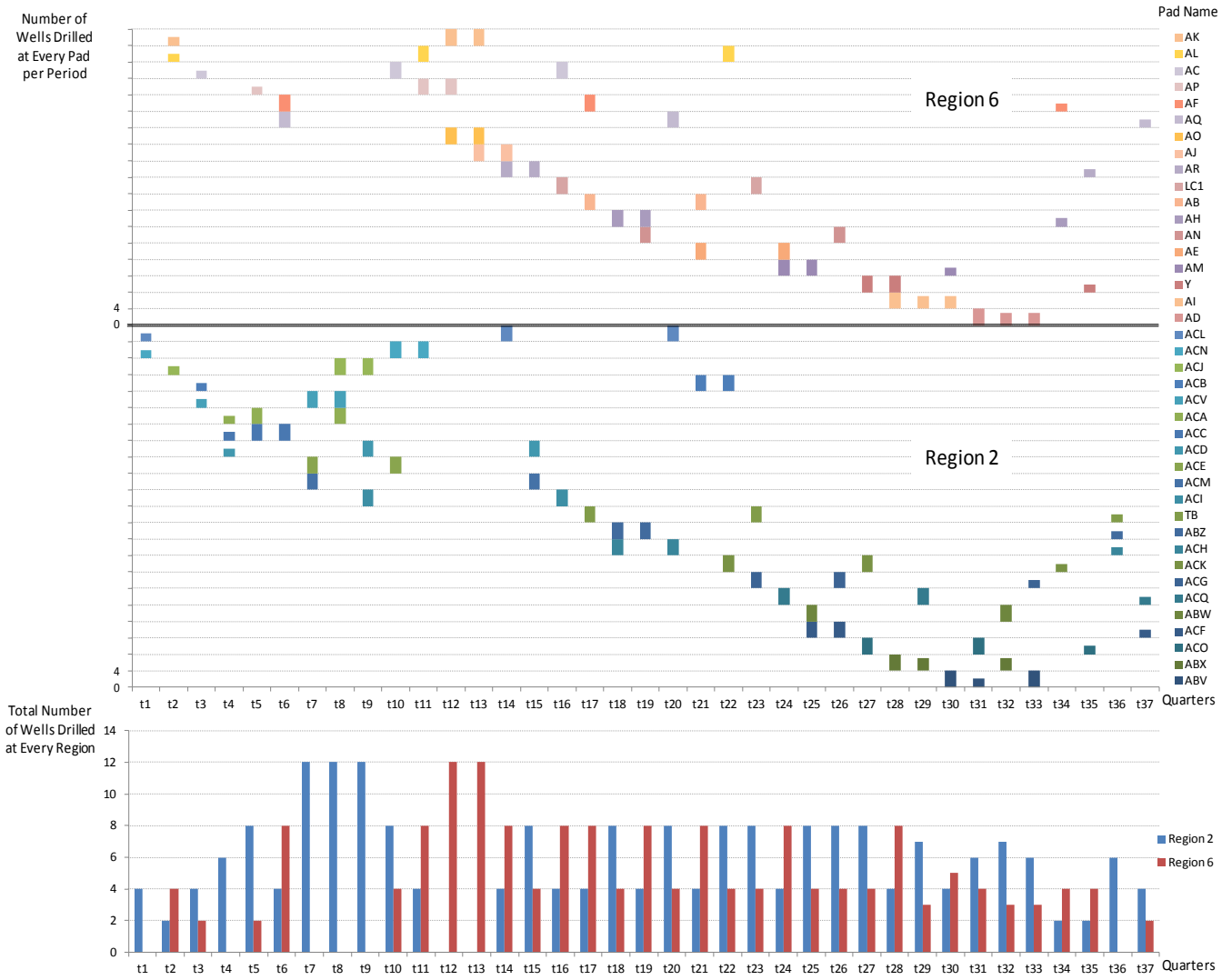

Figure 15. Optimal drilling strategy for the real-world case study.

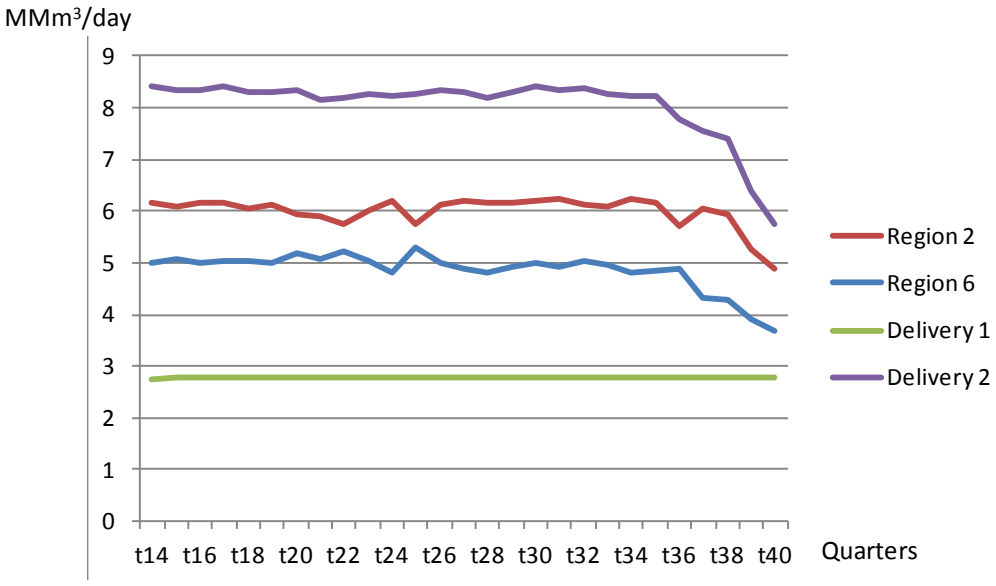

Figure 16. Shale gas flows in the optimal solution of the real-world case study.

6. CONCLUSIONS AND FURTHER WORK

A new MINLP model for the strategic planning of the shale gas supply chain has been presented in this work. The proposed formulation determines many of the critical decisions to be simultaneously optimized in the development of a shale gas project: the drilling and fracturing plan over time; the location, sizing and expansion of gas processing and fractionation plants; the section, length and location of gas and liquid pipelines (the network configuration); the power of gas compressors; and the amount of freshwater used for well drilling and fracturing, transported from available reservoirs; so as to maximize the economic results of the project. All the problem conditions such as flow balances, equipment sizing and expansions are modeled in linear constraints, while concave terms arise in the objective function due to the economies of scale determining the cost of plants, pipelines and compressors. Moreover, through a simple adaptation, the model can also account for shale gas composition variations depending on the geographic location of the wells.

Since the model becomes intractable for commercial global optimizers, a two-level decomposition algorithm successively refining piecewise linear approximations of the concave cost terms and solving reduced MINLP problems was implemented. The use of the δ-piecewise linear formulation[30] yields good relaxations of the MILP models, while the logarithmic approximation recently proposed by Cafaro and Grossmann[29] avoids numerical difficulties in the execution of the non-global MINLP solver. The proposed Branch-Reduce-Optimize (BRO) algorithm proves to be a useful tool for solving large-scale supply chain design problems in reasonable CPU times, although reducing the global optimality gap below 2.5% is quite hard for more challenging problems.

Results on realistic instances show the importance of heavier hydrocarbons to the economics of the project, and how the optimal planning of the drilling/fracturing strategy maximizes the utilization of gas processing/transportation infrastructure and improves the use of water resources. A real-world case study of the shale gas industry in north-western Pennsylvania involving more than 150 potential sites for well-pads was successfully solved. The solution obtained is of particular importance for industrial

decision-makers, who cannot readily optimize the drilling strategy together with the pipeline configuration and compressor sizing so as to obtain a higher profit.

Future work will focus on the optimization of the gas pipeline pressures, as well as the consideration of stochastic conditions for products demands, gas prices, water availability and shale gas production profiles at the wells.

ACKNOWLEDGMENTS

Financial support from Fulbright Commission Argentina, CONICET and CAPD at Carnegie Mellon University is gratefully acknowledged. We are also most grateful to EQT Corporation for the case study provided to us for Section 5.5.

NOTATION

Sets

F	Freshwater sources
I	Well-pads
J	Junction nodes
K	Gas demand points
P	Gas processing and fractionation plants
L	Ethane demand points
T	Time periods

Parameters

dr	Annual discount rate
$ethdem_{k,t}$	Maximum demand for ethane at node k in period t
$ethp_t$	Unit price of ethane in period t (forecast)
fix_f	Unit cost for freshwater acquisition from source f
$fwa_{f,t}$	Amount of freshwater available from source f during period t

$gasdem_{k,t}$ Maximum demand for methane at node k in period t

$gasp_t$ Unit price of methane in period t (forecast)

gc_i, ec_i, lc_i Methane, Ethane and LPG composition of the shale gas produced in pad i

ks, kd, kp, kc Base cost of plants, wells, pipelines and compressors in economy of scale functions

$l_{i,j}$ Distance between nodes i and j

$lpgp_t$ Average unit price of LPGs in period t (forecast)

$\overline{n}_i$ Upper bound on the number of wells to drill in pad i during one period

$\overline{N}_i$ Upper bound on the number of wells to drill in pad i over the planning horizon

$pw_{i,a}$ Daily shale gas production of a well of age a (periods) drilled in pad i

rf_i Water reuse factor in well-pad i

$sepmax$ Upper bound on the shale gas processing capacity of a single plant

s_g specific gravity in standard conditions

$shgp_t$ Unit cost of shale gas in period t

var_f Unit cost for freshwater transportation from source f

wr_i Amount of water required to drill and fracture a single well in pad i

τs, τp, τc Lead times for installing gas plants, pipelines and compressors, in quarters

Binary Variables

w_p = 1 if the processing plant p is operative during the planning horizon

$y_{i,n,t}$ = 1 if n wells are drilled at pad i during period t

Continuous Variables

$DFP_{i,j,t}$ Diameter of the gas pipeline installed between i and j in period t

$DGP_{j,p,t}$ Diameter of the gas pipeline installed between j and p in period t

$DTP_{p,k,t}$ Diameter of the gas pipeline installed between p and k in period t

$DLP_{p,l,t}$ Diameter of the liquid pipeline installed between p and l in period t

$EComp_{j,t}$ Ethane composition of the shale gas flow at the outlet of node j during period t

$FP_{i,j,t}$ Shale gas flow from well-pad i to junction j during period t

$FP^{E}_{i,,j,t}$ Individual ethane gas flow from well-pad i to junction j during period t

$FP^{G}_{i,,j,t}$ Individual methane flow from well-pad i to junction j during period t

$FP^{L}_{i,,j,t}$ Individual LPG flow from well-pad i to junction j during period t

$FPCap_{i,j,t}$ Total shale gas transportation capacity between i and j in period t

$FPFlow_{i,j,t}$ Shale gas transportation capacity installed between i and j in period t

$GComp_{j,t}$ Methane composition of the shale gas flow at the outlet of node j during period t

$GP_{j,p,t}$ Shale gas flow from junction node j to plant p during period t

$GPCap_{j,p,t}$ Total shale gas transportation capacity between j and p in period t

$GPFlow_{j,p,t}$ Shale gas transportation capacity installed between j and p in period t

$JCInst_{p,t}$ Compression power installed at node j in period t

$JCP_{p,t}$ Total compression power at j in period t

$LComp_{j,t}$ LPG composition of the shale gas flow at the outlet of node j during period t

$LP_{p,l,t}$ Ethane flow from plant p to demand point l during period t

$LPCap_{p,l,t}$ Total ethane transportation capacity between p and l in period t

$LPFlow_{p,l,t}$ Ethane transportation capacity installed between p and l in period t

$N_{i,t}$ Number of wells drilled in pad i during period t

$NP_{p,t}$ Daily production of LPG in plant p during period t

$PCInst_{p,t}$ Compression power installed at plant p in period t

$PCP_{p,t}$ Total compression power at p in period t

$TP_{p,k,t}$ Dry gas (methane) flow from plant p to demand point k during period t

$TPCap_{p,k,t}$ Total methane transportation capacity between p and k in period t

$TPFlow_{p,k,t}$ Methane transportation capacity installed between p and k in period t

$SepCap_{p,t}$ Total shale gas processing capacity of plant p in period t

$SepInst_{p,t}$ Daily shale gas processing capacity installed in plant p at period t

$SP_{i,t}$ Daily shale gas production of well-pad i during period t

$SP^{E}_{i,t}$ Daily ethane production of well-pad *i* during period *t*

$SP^{G}_{i,t}$ Daily methane production of well-pad *i* during period *t*

$SP^{L}_{i,t}$ Daily LPG production of well-pad *i* during period *t*

$WS_{f,i,t}$ Amount of freshwater supplied from source *f* to pad *i* during period *t*

APPENDIX A: Pipeline Flow, Compressor Power and Cost Calculations

Gas Pipeline Diameter, Flow and Cost

Similar to Durán and Grossmann,[8] the head loss in a gas pipeline segment *i-j* with diameter $D_{i,j}$ (in m), either transporting raw gas or methane, is assumed to be given by the Weymouth[21] flow equation (A1).

$$D_{i,j} = l_{i,j}^{\alpha} (P_i^2 - P_j^2)^{-\alpha} (B_{i,j})^{\alpha} \quad \text{(A1)}$$

where

$$B_{i,j} = s_g \, T [P_o / (0.375 T_o)]^2 \, (Flow_{i,j})^2 \quad \text{(A2)}$$

s_g is the gas specific gravity (0.729 kg/m^3 for shale gas, 0.554 kg/m^3 for methane) in standard conditions (P_o = 0.1013 MPa; T_o = 288.9 K). T is the average gas temperature, in this case fixed at 288.9 K, and α = 3/16. The input and output pressures (P_i, P_j, in MPa) are assumed to be known (see assumption 9 in Section 2) as well as the pipeline length $l_{i,j}$ (in km). By combining A1 and A2, the gas flow ($Flow_{i,j}$ in MMm^3/day) can be expressed by eq. (A3).

$$Flow_{i,j} = \{(P_i^2 - P_j^2) / s_g T [P_o / (0.375 T_o)]^2\}^{1/2} l_{i,j}^{-1/2} D_{i,j}^{1/2\alpha} \quad \text{(A3)}$$

As a result, if the inlet and outlet pressures are given, the gas flow is a function of the pipeline diameter to the power of 2.667, as shown in eq. (A4).

$$Flow_{i,j} = k_{i,j} \, l_{i,j}^{-0.5} D_{i,j}^{2.667} \quad \text{(A4)}$$

If shale gas pipelines transport raw gas from 2.1 to 1.4 MPa, the value of $k_{i,j}$ is 115.35. If the diameter is given in inches, it is 0.006423. On the other hand, dry gas pipelines operating from 6.0 MPa to 4.0 MPa show a value of $k_{i,j}$ equal to 378.06, or 0.02105 if the diameter is given in inches.

Finally, we use the economy of scale function (A5) to determine the cost of the gas pipeline *i-j*.

$$Cost_{i,j} = kpl_{i,j}\ l_{i,j}\ D_{i,j}^{\ 0.60} \tag{A5}$$

By substituting $D_{i,j}$ with the variable $DP_{i,j} = D_{i,j}^{2.667}$, eqs. (A4) and (A5) yield (A6) and (A7), which are the equations actually used in the MINLP model.

$$Flow_{i,j} = k_{i,j}\ l_{i,j}^{\ -0.5} DP_{i,j} \tag{A6}$$

$$Cost_{i,j} = kpl_{i,j}\ l_{i,j}\ DP_{i,j}^{\ 0.225} \tag{A7}$$

Liquid Pipeline Diameter, Flow and Cost

To calculate liquid flows in a pipeline *p-l*, a mean velocity (normally equal to $v^{\max}$ = 1.5 m/s) is assumed. That yields eq. (A8).

$$Flow_{p,l} = 86{,}400\ \pi / 4\ v^{\max}\ \rho\ D_{p,l}^{\ 2} \tag{A8}$$

where $Flow_{p,l}$ is given in ton/day, D (diameter) in meters, 86,400 is the total number of seconds per day the pipeline remains operative, and ρ is the liquid density, in ton/m^3 (0.546 ton/m^3 for liquid ethane).

Using the concave cost function given in (A5), and substituting $D_{p,l}$ with the variable $DP_{p,l} = D_{p,l}^{\ 2}$ yield eqs. (A9) and (A10), which are the equations finally used in the MINLP model.

$$Flow_{p,l} = k_{p,l}\ DP_{p,l} \tag{A9}$$

$$Cost_{p,l} = kpl_{p,l}\ l_{p,l}\ DP_{i,j}^{\ 0.3} \tag{A10}$$

Compression Power

Since the compressors are assumed to be adiabatic, the power requirement of a compressor installed at node j (CP_j in kW) can be calculated through eq. (A11).[8]

$$F_j \; CP_j = (Pd_j / Ps_j)^b - 1 \tag{A11}$$

where

$$F_j = [(\gamma - 1)\, \eta / (4.0426\, T\, \gamma)]\, Flow_j^{-1} \tag{A12}$$

$$b = z\, (\gamma - 1) / \gamma \tag{A13}$$

z is the gas compressibility factor (by the ideal gas assumption, z = 1), γ is the heat capacity ratio (typically, γ = 1.26), η is the compressor efficiency, and T is the gas temperature at suction conditions (T = 288.9 K). $Flow_j$ is given in MMm^3/day. By assumption 9 (see Section 2) the compression ratio (Pd_j / Ps_j) both at junction and plant compressors is given (usually equal to 1.5). Hence, combining (A11), (A12) and (A13) yields eq. (A14), stating that the power requirement is linearly proportional to the gas flow (see eqs. (27) and (28) of the MINLP model).

$$CP_j = [(4.0426\, T\, \gamma) / (\gamma - 1)\, \eta][(Pd_j / Ps_j)^{z(\gamma-1)/\gamma} - 1]\, Flow_j = kc_j \; Flow_j \tag{A14}$$

APPENDIX B: Other Model Features

Delayed Production of a Well

Some companies may often drill, fracture and complete a non-conventional gas well, but the production of the well is delayed until the required infrastructure (pipelines, compressors, etc.) becomes available. In that case, the model is adapted by incorporating an integer variable accounting for the number of wells of pad i that become productive at the beginning of period t ($NP_{i,t}$). Then eq. (B1) is added to the formulation, and eq. (4) in the original model is replaced by (B2).

$$\sum_{\tau \le t} NP_{i,\tau} \le \sum_{\tau < t} N_{i,\tau} \qquad \forall i \in I, t \in T \tag{B1}$$

$$\sum_{\tau=1}^{t} NP_{i,\tau}\ pw_{i,t-\tau+1} = SP_{i,t} \qquad \forall i \in I, t > 1 \tag{B2}$$

Cost of Rigs and Crew for Drilling New Wells

If the cost of moving rigs, drilling crews and other resources from one pad to the other is significant, the model should be able to determine the period in which the crew arrives at a pad to start or continue drilling new wells. With that purpose, we incorporate a new binary variable $x_{i,t}$ that is equal to one if at least one well is drilled and fractured in pad i during period t. That is controlled by eqs. (B3) and (B4).

$$N_{i,t} \le \overline{N}_i\ x_{i,t} \qquad \forall i \in I, t \in T \tag{B3}$$

$$N_{i,t} \ge x_{i,t} \qquad \forall i \in I, t \in T \tag{B4}$$

As a result, the cost of arriving at a well pad i to start or continue the drilling of new wells in period t is lower bounded by eq. (B5), and is included in the objective function (35).

$$RC_{i,t} \ge rigc_i (x_{i,t} - x_{i,t-1}) \qquad \forall i \in I, t \in T \tag{B5}$$

Finally, if the total number of rigs (and/or drilling crews) available is *rigmax*, eq. (B6) imposes an upper bound on the number of pads where new wells are drilled during a single period.

$$\sum_{i \in I} x_{i,t} \le rigmax \qquad \forall t \in T \tag{B6}$$

LITERATURE CITED

[1] U.S. Energy Information Administration (EIA). *U.S. 2012 Annual Energy Outlook with Projects to 2035*. Washington, DC: US Department of Energy, 2012.

[2] U.S. Energy Information Administration (EIA). *Natural Gas Processing Plants in the United States: 2010 Update*. Washington, DC: US Department of Energy, 2011.

[3] Laiglecia JI, Lopez-Negrete R, Diaz MS, Biegler LT. A simultaneous dynamic optimization approach for natural gas processing plants. *Proceedings of Foundations of Computer Aided Process Operations (FOCAPO)*, 2012.

[4] Ladlee J, Jacquet J. *The Implications of Multi-Well Pads in the Marcellus Shale*. *Research & Policy Brief Series.* Ithaca, NY: Cornell University's Community & Regional Development Institute (CaRDI), 2011.

[5] Stark M, Allingham R, Calder J, Lennartz-Walker T, Wai K, Thompson P, Zhao S. *Water and Shale Gas Development. Leveraging the US experience in new shale developments*. Dublin, Ireland: Accenture, 2012.

[6] Geoffrion AM, Graves GW. Multicommodity distribution system design by Benders decomposition, *Management Science*. 1974; 20: 822-844.

[7] Melo MT, Nickel S, Saldanha-da-Gama F. Facility location and supply chain management – A review. *European Journal of Operational Research.* 2009; 196: 401-412.

[8] Durán MA, Grossmann I.E. A mixed-integer nonlinear programming algorithm for process systems synthesis. *AIChE J.* 1986; 32: 592-606.

[9] Iyer RR, Grossmann IE. Vasantharajan S, Cullick AS. Optimal planning and scheduling of offshore oil field infrastructure investment and operations. *Ind. Eng. Chem. Res.* 1998; 37: 1380-1397.

[10] Van Den Heever SA, Grossmann IE. An iterative aggregation/disaggregation approach for the solution of a mixed-integer nonlinear oilfield infrastructure planning model. *Ind. Eng. Chem. Res.* 2000; 39: 1955-1971.

[11] Gupta V, Grossmann IE. An efficient multiperiod MINLP model for optimal planning of offshore oil and gas field infrastructure. *Ind. Eng. Chem. Res*. 2012; 51: 6823-6840.

[12] Rahman MM, Rahman MK, Rahman SS. An integrated model for multiobjective design optimization of hydraulic fracturing. *Journal of Petroleum Science and Engineering*. 2001; 31: 41-62.

[13] Knudsen BR, Foss B, Grossmann IE, Gupta V. Lagrangian relaxation based production optimization of tight-formation wells. . Submitted for publication to *Comput. Chem. Eng.* 2013.

[14] Mauter MS, Palmer VR, Tang Y, Behrer RP. *The next frontier in United States unconventional shale gas and tight oil extraction: Strategic reduction of environmental impact*. Belfer Center for Science and International Affairs, Harvard Kennedy School, March 2013.

[15] Rahm BG, Riha SJ. Toward strategic management of shale gas development: Regional, collective impacts on water resources. *Environmental Science and Policy*. 2012; 17: 12-23.

[16] Yang L, Grossmann IE. Superstructure-Based Shale Play Water Management Optimization. 2013 AIChE Meeting. San Francisco, CA.

[17] Nikolaou M. Computer-aided process engineering in oil and gas production. *Comput. Chem. Eng.* 2013; 51: 96-101.

[18] Troner A. *Natural gas liquids in the shale revolution*. James A. Baker III Institute for Public Policy Rice University. April, 2013.

[19] Grossmann IE. Enterprise-wide optimization: A new frontier in process systems engineering. *AIChE J*. 2005; 51: 1846-1857.

[20] Oliveira F, Gupta V, Hamachera S, Grossmann IE. A Lagrangean decomposition approach for oil supply chain investment planning under uncertainty with risk considerations. *Comput. Chem. Eng.* 2013; 50: 184-195.

[21] Weymouth TR. Problems in natural gas engineering. *ASME Transactions.*1942; 34:185-234.

[22] Nahmias, S. *Production and operations analysis*. New York, NY: McGraw-Hill, 2009.

[23] Biegler LT, Grossmann IE, Westerberg AW. *Systematic methods of chemical process design*. New Jersey, NJ: Prentice Hall, 1997.

[24] Guthrie KM. Capital cost estimating. *Chemical Engineer*. 1969; 76: 114-142.

[25] McCarl BA. *Expanded GAMS user guide version 23.6*. Washington, DC: GAMS Development Corporation, 2011.

[26] Geoffrion AM. Objective function approximations in mathematical programming. *Math. Prog.* 1977; 13: 23-37.

[27] Bergamini ML, Aguirre P, Grossmann IE. Logic-based outer approximation for globally optimal synthesis of process networks. *Comput. Chem. Eng.* 2005; 29: 1914-1933.

[28] Bergamini ML, Grossmann IE, Scenna N, Aguirre P. An improved piecewise outer-approximation algorithm for the global optimization of MINLP models involving concave and bilinear terms. *Comput. Chem. Eng.* 2008; 32: 477-493.

[29] Cafaro DC, Grossmann IE. Alternate approximation of concave cost functions for process design and supply chain optimization problems. Submitted for publication to *Comput. Chem. Eng.* 2013.

[30] Padberg M. Approximating separable nonlinear functions via mixed zero-one programs. *Oper. Res. Lett.* 2000; 27: 1-5.

[31] You F, Grossmann IE. Stochastic inventory management for tactical process planning under uncertainties: MINLP Model and Algorithms. *AIChE J*. 2011; 57: 1250-1277.

[32] You F, Grossmann IE. Integrated multi-echelon supply chain design with inventories under uncertainty: MINLP models, computational strategies. *AIChE J.* 2010; 56: 419-440.