Journal : SmallExtended 11750 | Article No : 635 | Pages : 30 | MS Code : 635 | Dispatch : 13-6-2022

TOP
https://doi.org/10.1007/s11750-022-00635-3

INVITED PAPER

Recent contributions to the optimal design of pipeline networks in the energy industry using mathematical programming

Diego C. Cafaro[1,2] · Demian J. Presser[1,2] · Ignacio Grossmann[3]

Received: 30 December 2021 / Accepted: 25 May 2022

Abstract
The optimal design of pipeline networks has inspired process systems engineers and operations research practitioners since the earliest times of mathematical programming. The nonlinear equations governing pressure drops, energy consumption and capital investments have motivated nonlinear programming (NLP) approaches and solution techniques, as well as mixed-integer nonlinear programming (MINLP) formulations and decomposition strategies. In this overview paper, we present a systematic description of the mathematical models proposed in recent years for the optimal design of pipeline networks in the energy industry. We provide a general framework to address these problems based on both the topology of the network to build, and the physical properties of the fluids to transport. We illustrate the computational challenges through several examples from industry collaboration projects, published in recent papers from our research group.

Keywords Pipeline · Network · Energy · Supply chain · Design · Optimization · MINLP

List of Symbols

Sets

A_i	Subset of nodes adjacent to node i
C	Components in the fluid stream
D	Alternative pipeline diameters

Diego C. Cafaro
dcafaro@fiq.unl.edu.ar

1 INTEC(UNL-CONICET), Güemes 3450, 3000 Santa Fe, Argentina

2 Fac. de Ing. Química, Universidad Nacional del Litoral, Santiago del Estero 2829, 3000 Santa Fe, Argentina

3 Department of Chemical Engineering, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, USA

Journal : SmallExtended 11750 | Article No : 635 | Pages : 30 | MS Code : 635 | Dispatch : 13-6-2022

I,J,K,L	Nodes in the network
S	Fluid states
T	Time periods

Parameters

ec_t	Unit energy cost during period t
g	Gravity constant
glr	Gas-to-liquid ratio in multiphase flows
$k_{i,j}$	Constant of the Hazen-William correlation for a pipeline connecting i and j
$l_{i,j}$	Length of the pipeline connecting i and j
lt	Lead-time for pipeline construction
r	Discount rate for cashflows
$\mathrm{sg},\ \mathrm{sl}$	Specific gravity of the gas/liquid
s_{L}	Head loss per unit length for water pipeline design purposes
$P_{\mathrm{o}},\ T_{\mathrm{o}}$	Pressure and temperature at standard conditions
$\mathrm{tc}_{i,j,d}$	Transportation capacity of a pipeline with diameter d connecting i and j
$\mathrm{vmax}_{\mathrm{LP/GP}}$	Maximum linear velocities admitted for liquid and gas phases
z	Gas compressibility factor
α, β	Parameters of economy-of-scale functions
γ	Constant of the Weymouth correlation for gas flows
$\delta_{i,j,d}$	Numerical value of the diameter d for a pipeline connecting i and j
$\Delta\mathrm{sp}_{i,j}^{\mathrm{Max}}$	Maximum difference of square pressures allowed between nodes i and j
Δt	Length of a time period
$\Delta z_{i,j}$	Elevation difference between nodes i and j
ε	Roughness of the internal wall of the pipeline
ζ	SPE constant for multiphase pipeline sizing
η	Pump yield
θ	Temperature of the fluid
ν	Kinematic viscosity of the fluid
μ_{c}	Relative contribution of component c to the calculation of the pressure drop
$\rho, \rho_{\mathrm{avg}}$	Density/average density of the multiphase flow
$\varphi_{c,j}$	Fraction of component c removed from the flow stream at node j
$\psi_{s,s',i}$	Yield of s' per unit of s processed in node i
ω	Exponent of the Hazen-Williams correlation

Non-negative variables

$\mathrm{Capex}\left(d_{i,J,t}\right)$	Capital expenditures on a pipeline with diameter d connecting i and j built in period t
$d_{i,J,t}$	Diameter of a pipeline connecting i and j built in period t

Journal : SmallExtended 11750 | Article No : 635 | Pages : 30 | MS Code : 635 | Dispatch : 13-6-2022

$dl_{i,J,t}$	$= d_{i,j,t}^{2}$ for liquid pipelines
$dg_{i,J,t}$	$= d_{i,j,t}^{2.667}$ for gas pipelines
$D_{c,K,t}$	Demand of component *c* at node *k* during period *t*
$D_{s,I,t}$	Demand of fluid in state *s* at node *i* during period *t*
$C_{s,I,t}$	Amount of fluid in state *s* processed at node *i* during period *t*
f	Friction factor
$F_{c,I,j,t}$	Amount of component *c* flowing from *i* to *j* during period *t*
$F_{s,I,i',t}$	Amount of fluid in state *s* flowing from *i* to *i’* during period *t*
h_{L}	Head loss due to friction
$I_{s,I,t}$	Inventory level of *s* at node *i* at the end of period *t*
NPC	Total net present cost
$\mathrm{Opex}\left(F_{c,I,j,t}\right)$	Operating expenditures of a pipeline carrying *F* units of *c* from *i* to *j* during period *t*
$P_{i,T}$	Pressure at node *i* during period *t*
$P_{j,T}^{\mathrm{sq}}$	$P_{i,t}^{2}$ for gas and multiphase pipelines
PW_{L}	Pump power required to compensate head loss due to friction
$Q_{c,I,t}$	Production of component *c* at node *i* during period *t*
$Q_{s,I,t}$	Production of fluid in state *s* at node *i* during period *t*
$R_{s,I,t}$	Amount of fluid converted into state *s* at node *i* during period *t*
Re	Reynolds number
$\mathrm{TC}\left(d, P_{i,t}, P_{j,t}\right)_{i,j,\tau}$	Transportation capacity of a pipeline with diameter *d* connecting *i* and *j* built in period *τ*, according to inlet and outlet pressures in period *t*
$\mathrm{TC}_{i,J,t}^{\mathrm{oil/gas/water}}$	Transportation capacity of oil/gas/water through pipeline *i*–*j* built in period *t*
U	Mean linear velocity of the flow
$X_{c,J,t}$	Concentration of *c* in the flows leaving node *j* during period *t*
X_{LM}	Lockhart-Martinelli parameter for multiphase flows
$\Delta P_{i,J,t}^{\mathrm{sq}}$	Difference of square pressures at nodes *i* and *j* during period *t*
$\Delta Pg,\ \Delta Pl$	Pressure drops for gas and liquid phases in multiphase flows

Binary variables

$u_{i,J,t}$	= 1 if the flow direction is from *i* to *j* during period *t*
$x_{i,J,d,t}$	= 1 if a pipeline with diameter *d* is built between *i* and *j* in period *t*

1 Introduction

The energy industry is currently facing difficult challenges due to increasing competitiveness and narrow profit margins. Besides that, and perhaps more importantly, the need to adopt cleaner sources that reduce the impact of carbon emissions and

Journal : SmallExtended 11750	Article No : 635	Pages : 30	MS Code : 635	Dispatch : 13-6-2022

freshwater consumption are also unavoidable. In this context, one of the relevant problems that have received increasing attention is the optimal design and management of pipeline networks. Pipeline network design aims to optimally determine the connections (or links) between a set of nodes, the location of junction points and the diameters of the pipelines (Bhaskaran and Salzborn 1979), while network management is geared towards the operational planning of pipes and facilities to meet demands under customer specifications and contractual rules (Selot et al. 2008).

Pipelines have been widely used in the fossil energy industry since 1862. Modern oil and gas (O&G) supply chains comprise extensive networks of pipelines that may carry oil, oil refined products, natural gas, natural gas liquids (NGL), liquefied natural gas (LNG) and methanol, among other energy carriers. In addition, the O&G industry has recently addressed the design of pipeline networks carrying other fluids in large scales for the production of energy. There are two particular problems of current importance in the O&G sector: the supply of freshwater plus the gathering, processing and recycle of produced water to/from unconventional wells for hydraulic stimulation (Cafaro and Grossmann 2020); and the capture, storage, distribution and use of carbon dioxide, mainly to enhance oil recovery from mature fields (Presser et al. 2022). Moreover, the energy industry is looking for cleaner energy carriers that yield lower impacts to the environment but still have a high energy density. That is the case of hydrogen (in liquid or gas state), and related chemicals like ammonia (an easy-handling option in the hydrogen supply chain).

Given the benefits of pipelines as transportation means from economic, environmental and reliability standpoints (Cafaro and Cerdá 2012), designing efficient networks of pipelines to aid the sustainable production and supply of energy products is a relevant problem. The main reason why pipeline design decisions are so critical is that they usually imply very large investment costs (in the order of millions of dollars per kilometer of length), and there are substantial benefits from the economies of scale. Thus, sizing pipelines and managing operations in a way that these networks are efficiently utilized is the key to the economic success of many energy projects.

According to Mah and Shacham (1978), a pipeline network is a collection of elements such as pipes, compressors, pumps, valves, regulators, heaters, tanks, and reservoirs interconnected in a specific way. In their early review on the optimal design of pipeline networks, they distinguish among two parts of the problem that govern the network performance. They can be synthesized as topological and fluid dynamics considerations. The first characteristic refers to the way to connect the elements, while the second is determined by physical laws. From the mathematical programming perspective, the optimal topology of a pipeline network can be addressed by means of integer programming models, in which the main decisions are represented by 0–1 variables accounting for the installation of a pipeline segment between a pair of the nodes. Pipeline interconnections are usually selected from a superstructure of alternatives (Montagna et al. 2021). Material flows (continuous variables) and balances (linear constraints) are also included in this part of the problem, leading to mixed-integer linear programming (MILP) formulations. These can be viewed as particular cases of the network flow problem, closely related to electrical circuit design and analysis.

Besides the combinatorial complexity of the topological problem, the need to simultaneously size the pipelines with the selection of the optimal diameter for every connection makes it even more difficult. Fluid dynamics laws are highly nonlinear, both for liquid and gas streams, yielding complex relations between mass flowrates and pressure losses, according to the pipeline length and diameter. As a result, and generally speaking, the optimal design and operation of pipeline networks leads to mixed-integer nonlinear programming (MINLP) formulations that are computationally challenging to solve. That is why as suggested by Mah and Shacham in (1978), not all the questions raised in the pipeline network design and synthesis can be completely and satisfactorily answered. However, great progress has been made in the last decade as it is shown in this contribution.

The aim of this paper is to summarize recent contributions and provide a generalized, systematic framework to address pipeline network design problems for the energy industry using mathematical programming. Most of the models described in this work are the results of recent collaboration projects jointly developed at the Center for Advanced Process Decision-Making (CAPD) of the Carnegie Mellon University (Pittsburgh, PA, USA) and the INTEC (University of Litoral and CONICET, Santa Fe, Argentina). The projects have been particularly focused in the optimal design and operation of pipeline networks, in close association with industry partners.

2 General formulations for pipeline network design

Designing a network of pipelines to transport gas and liquids in the energy industry generally comprises three main decisions: (a) how to connect the nodes (pipeline layout), (b) what is the pipeline diameter required for each segment, and (c) what are the flows moving along each pipeline segment during the time horizon. Moreover, as highlighted by Saldanha-da-Gama (2018), the location of structures in a real-world network may require including time as an extra dimension in the decision making process. Indeed, for planning purposes, the time domain is typically discretized in days, weeks, months or quarters. The geographical location of producing sources, potential intermediate nodes and demand sites are usually given, from which the length of each possible connection is known beforehand. However, building a pipeline to connect each pair of nodes is a decision to be determined. Depending on the number of segments along which the fluids move to reach the destination nodes, one may distinguish between two different types of formulations to optimize the pipeline network design: (1) mathematical models with a fixed number of steps or "echelons", and (2) formulations with an undetermined number of echelons.

2.1 Formulations with a fixed number of echelons

These formulations assume that there is a set of nodes i (producers) that need to be connected to one or several second level nodes j (junctions), which in turn need to be connected to one or many third level nodes (e.g., processing nodes), until reaching the demand nodes. A simple example for gathering and processing shale gas

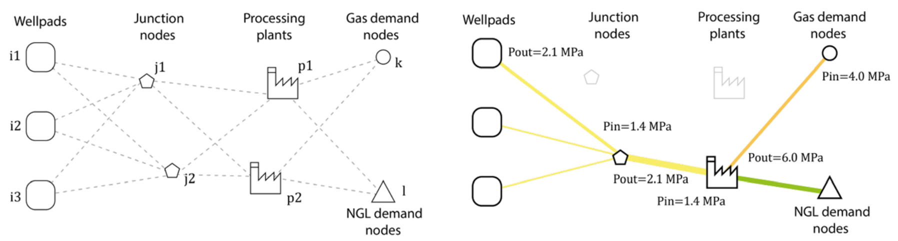

Fig. 1 Simplified superstructure of alternatives (left) and a possible network design in the shale gas supply chain (right), assuming a predefined number of echelons. Pressures are given in 10^6 Pascal (MPa)

is presented in Fig. 1. While the left side of Fig. 1 shows all possible connections between the nodes in the network, the right side shows a feasible solution for this illustrative problem. It is interesting to note that the product state evolves with each stage. In the example of Fig. 1, the raw gas produced in nodes i is dehydrated and compressed in junction nodes j, separated in processing nodes p, and the components (natural gas, ethane and LPG) are finally consumed at demand nodes k and l.

There are basically two blocks of equations in this formulation, namely mass balances and pipeline sizing constraints. Mass balances are imposed at every intermediate node (generically denoted by j) and they state that the summation of all inlet flows (coming from nodes of type i) equals the overall outlet flows (connecting to nodes k). That is represented by Eq. (1), in which we assume, for simplicity, that storing material in the intermediate nodes is not an option. In the most general case, the streams comprise an overall flow (e.g., natural gas) that is a mixture of individual components (e.g., methane, ethane, propane, etc.). The set C accounts for every single component, also including the element $c_o \in C$ that represents the overall stream. The parameter $\varphi_{c,j}$ is usually used when the node j is a processing facility where components in the flow are separated from the mainstream. For instance, in a gas dehydrating unit located at j, $\varphi_{c_o,j} < 1$ since a component in the overall flow (water) is removed from the stream.

$$\varphi_{c,j} \sum_{i} F_{c,i,j,t} = \sum_{k} F_{c,j,k,t} \quad \forall j \in J, c \in C, t \in T. \tag{1}$$

If the node j is a splitting node, i.e., fractions of the incoming flow are directed to two or more destinations, extra considerations should be made to manage the mixture of components. To ensure a homogeneous distribution of the component flows that mix at node j, the bilinear Eq. (2) is included in the formulation.

$$F_{c,j,k,t} = X_{c,j,t} F_{c_o,j,k,t} \quad \forall j \in J, k \in K, c \in C - \{c_o\}, t \in T. \tag{2}$$

Variable $X_{c,j,t}$ is the concentration of c per unit volume of fluid pooled at node j during time period t. Since the index c_o accounts for the overall stream (i.e., including all components), the right-hand side of Eq. (2) yields the flowrate of the individual component c from j to k during period t. Hellemo and Tomasgard (2015) present a generalized MINLP formulation for the pooling problem, where flows of different components, subproducts and products are allocated to intermediates points in a pipeline network before being delivered to terminal points under specific quality constraints. Note that Eq. (2) leads to a nonconvex optimization problem that makes it difficult to search for the global optimal solution. This equation is usually avoided when: (a) the fluid being transported is not a mixture of components, (b) there is no need to track individual components but simply the overall flow, or (c) all the flows gathered at j need to converge to a single node k (no splitting assumption).

The second block of constraints deals with the pipeline sizing and installation. In general terms, if the model selects the link i–j at period t, that pipeline segment must be available by that time. Moreover, the magnitude of the flow should be less or equal to the pipeline transportation capacity between nodes i and j, as imposed by constraint (3). Note that this transportation capacity is a function of the pipeline diameter d, the fluid state and the pressures at the inlet and outlet extremes of the segment $(P_{i,t}, P_{j,t})$. The way to calculate the transportation capacity of a pipeline installed in period τ, namely $TC\left(d, P_{i,t}, P_{i',t}\right)_{s,i,i',\tau}$, is addressed in detail in upcoming sections.

$$\sum_{c} \mu_c F_{c,i,j,t} \leq \sum_{\tau \leq t-lt} TC\left(d, P_{i,t}, P_{j,t}\right)_{i,j,\tau} \quad \forall i \in I, j \in J, c \in C, t \in T. \tag{3}$$

The value of the parameter μ_c is a modeler's choice, based on the procedure followed to calculate the transportation capacity. For instance, in most natural gas gathering networks,$\mu_{c_o} = 1$ and $\mu_c = 0$ for all $c \neq c_o$, meaning that the transportation capacity simply restricts the overall flow. Instead, in oil gathering networks, $\mu_{\text{oil}} = 1$ while $\mu_{\text{gas}} = \mu_{\text{water}} = 0$, meaning that the transportation capacity is given in terms of the oil flowrate (Montagna et al. 2021), even though the flow is also comprised by water and gas phases. We note that the latter assumption relies on the condition of constant flow compositions. In Eq. (3) lt is the lead time to build the pipeline segment, given as an integer number of time periods.

Finally, production and demand constraints are imposed at the sources i and destination nodes k, respectively, as expressed through Eqs. (4) and (5). Depending on the problem goal, only one of these constraints is usually binding at the optimum. Furthermore, in problems in which the network design is integrated to the development planning, the production rate Q and/or the demand rate D are decision variables to be optimized. Since this work is focused on the design of pipeline networks, we refer the reader to the contributions by Cafaro and Grossmann (2014a) and Cafaro and Grosssmann (2020) for further details on how to model the planning of tasks like drilling, stimulation, completion and production, from which demand and/or production rates are obtained.

$$\sum_{j} F_{c,i,j,t} = Q_{c,i,t} \quad \forall i \in I, c \in C, t \in T, \tag{4}$$

$$\sum_{j} F_{c,j,k,t} = D_{c,k,t} \quad \forall k \in K, c \in C, t \in T. \tag{5}$$

If the formulation is strictly concerned with the pipeline network design, the objective function usually comprises capital and operating costs. These are brought to the present time to obtain the minimum net present cost. A general form of that function is given by Eq. (6), where r is the discount rate. Pipeline capital costs (Capex) are usually dependent on the diameter $d_{i,j,t}$ being selected to connect nodes i and j (built in period t), while operating costs (Opex) can be calculated from the material flows $F_{c,i,j,t}$ carried at every single period. Note that Capex and Opex are non-negative variables defined by model decisions.

$$\text{Min NPC} = \sum_{t} (1+r)^{-t} \left[\sum_{i,j} \text{Capex}\left(d_{i,j,t}\right) + \sum_{i,j,c} \text{Opex}\left(F_{c,i,j,t}\right) \right]. \tag{6}$$

The way to calculate variables Capex and Opex also depends on the nature of the problem. In Sect. 3, we present alternative formulations for different fluids (liquid, gas, multiphase), under different assumptions (discrete or continuous values for pipeline diameters, known or unknown pressures at inlet and outlet sections, etc.). A general MINLP formulation for pipeline network design, given a fixed number of echelons to connect production with consumption nodes, will seek to minimize function (6) subject to constraints (1) to (5).

2.2 Formulations with an undetermined number of echelons

These formulations can be regarded as generalizations of the previous approaches. In this case we assume that there is a set of nodes i (either sources, intermediate and/or demand nodes) that need to be connected to one or more nodes i' in the same set, which in turn need to be connected to nodes i'', until reaching the final destination nodes (see Fig. 2). In any of these nodes, facilities for merging, splitting, storing, separating and/or processing flows should be installed to make the product flows be ready for delivery or use. One of the major differences with regards to the models with a fixed number of echelons is that the flow direction may be reversed in any pipeline segment over the time horizon.

Similar to the models with a fixed number of echelons, mass balance equations are key constraints of these formulations. Equation (7) states that the overall inventory level at a certain node can be tracked over time by adding flows $F_{s,i',i,t}$ coming from adjacent nodes $i' \in A_i$, removing the streams $F_{s,i,i'',t}$ derived to other locations i'', subtracting the amounts $D_{s,i,t}$ and $C_{s,i,t}$ consumed in the same node i, and adding the amounts $R_{s,i,t}$ and $Q_{s,i,t}$ produced in that node during the same time period t. In this case, we model the equations considering material storage, but for simplicity, we omit the index c of individual components. Model extensions to track the flow composition are straightforward but might require including bilinear constraints, as explained in the previous section (for more details, see Cafaro and Grossmann 2020).

$$I_{s,i,t} = I_{s,i,t-1} + \sum_{i' \in A_i} F_{s,i',i,t} - \sum_{i'' \in A_i} F_{s,i,i'',t} - D_{s,i,t} - C_{s,i,t} + Q_{s,i,t} + R_{s,i,t} \quad \forall s \in S, i \in I, t \in T. \tag{7}$$

Journal : SmallExtended 11750 | Article No : 635 | Pages : 30 | MS Code : 635 | Dispatch : 13-6-2022

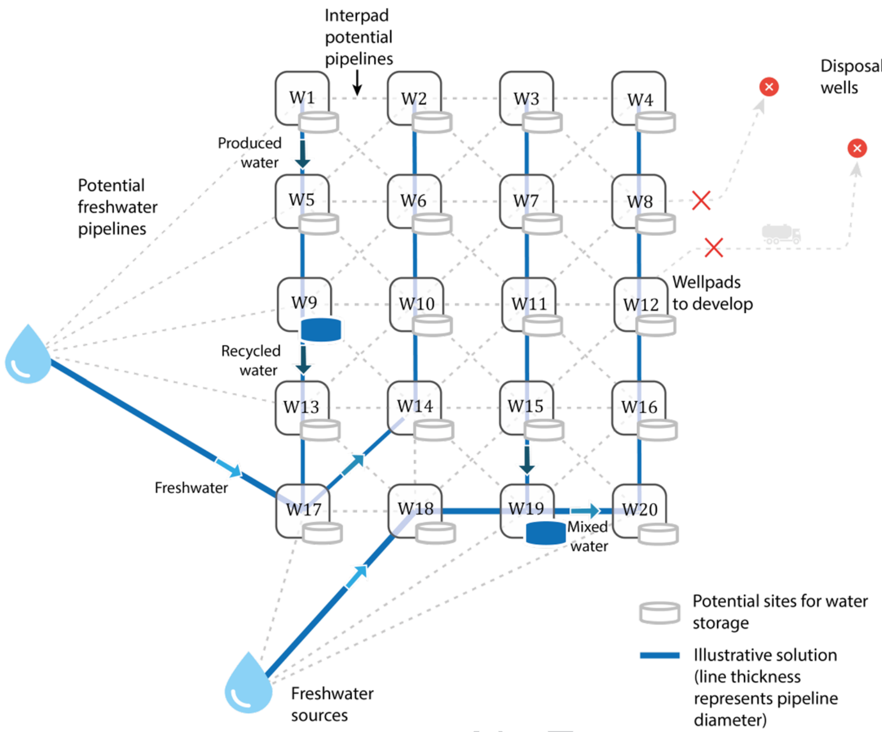

Fig. 2 Superstructure of alternatives for the design of water pipeline networks over a shale gas development area, with no predefined number of echelons

Although the indices for individual components are omitted, Eq. (7) includes the additional index s accounting for product states. If a processing facility is installed at node i, the consumption of material in state s (e.g., raw gas) may be associated with the production of s' (e.g., dry gas) in the same node, as denoted by Eq. (8), where $\psi_{i,s,s'}$ is the yield of s' per unit of s processed in node i. Figure 3 conceptually illustrates the links between the different product states (layers in Fig. 3) and their corresponding pipeline networks within the proposed model structure. Note that layers do not necessarily correspond to different installation depths of the pipelines below the ground.

$$R_{s',i,t} = \sum_{s} \psi_{s,s',i} C_{s,i,t} \quad \forall s' \in S, i \in I, t \in T. \tag{8}$$

Note that in Eq. (7) there are two terms accounting for the production of material in state s, which are represented by the variables $Q_{s,i,t}$ and $R_{s,i,t}$. In the first case, the additional material flow comes from external sources (e.g., well production); while in the second case, it results from processing other states in the same

Journal : SmallExtended 11750 | Article No : 635 | Pages : 30 | MS Code : 635 | Dispatch : 13-6-2022

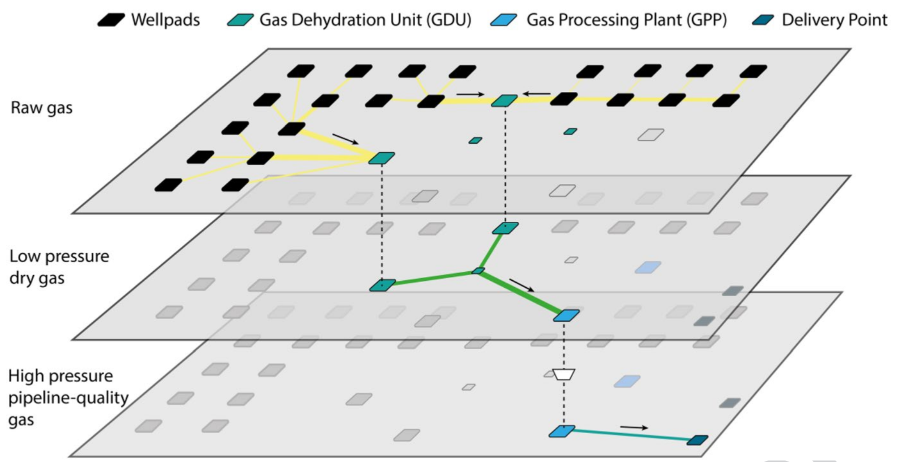

Fig. 3 Conceptual Illustrative example of a natural gas gathering network presenting one layer for each state (raw gas, low pressure dry gas, high pressure pipeline-quality gas) and its corresponding pipeline network

node i. In the most general case, external production $Q_{s,i,t}$ could be also dependent on the development planning.

Similarly, product consumption may be due to external demand $D_{s,i,t}$ or due to processing (change of state), which is represented by $C_{s,i,t}$. In summary, optimization models with no determined number of echelons yield "s" different layers of the pipeline network, one for each product state. Links between these networks are the processing facilities at the nodes, where the state of the products may change (for more details, see Montagna et al. 2022).

The second block of constraints accounts for pipeline sizing, imposing an upper bound on the flow of product in state s that is directed from i to i' during period t. More details on how to calculate the transportation capacity $TC\left(d, P_{i,t}, P_{i',t}\right)_{s,i,i',\tau}$ according to the pipeline diameter, pressures and product state are given in later sections.

$$F_{s,i,i',t} \leq \sum_{\tau \leq t-lt} TC\left(d, P_{i,t}, P_{i',t}\right)_{s,i,i',\tau} \quad \forall i \in I, i' \in A_i, s \in C, t \in T. \tag{9}$$

Although the model, under these constraints, does not strictly avoid bidirectional flows during the same time period (i.e., $F_{s,i,i',t} > 0$ while $F_{s,i',i,t} > 0$), this is discouraged by the operating costs term in the objective function (10) to be minimized. More specifically, one will unnecessarily pay for the transportation of the same material in both directions, yielding the same results in the mass balances. In particular cases, however, binary variables are required to explicitly determine the flow directions. This is also described in further sections.

$$\text{Min NPC} = \sum_{t} (1+r)^{-t} \left[\sum_{i<i'} \text{Capex}\left(d_{i,i',t}\right) + \sum_{i,i',s} \text{Opex}\left(F_{s,i,i',t}\right) \right]. \tag{10}$$

Finally, the general MINLP formulation for pipeline network design with no predefined number of echelons seeks to minimize function (10) subject to constraints (6)–(9).

3 Calculations of the maximum admissible flows, capital and operating costs

From the previous formulations, questions that still remain open are the following: (1) how to estimate the transportation capacity of the pipelines according to diameters and pressures, and (2) how to calculate capital and operating costs. In this section, we describe mathematical models to address these terms, distinguishing between liquid pipelines, gas pipelines, and multiphase (liquid + gas) pipelines.

3.1 Liquid pipeline networks

One of the most widely studied problems involving liquid pipelines is the optimization of water distribution networks where water demand rates at different nodes are given parameters (Caballero and Ravagnani 2019). However, there are problems for which water is a dependent demand item because its requirement is driven by the operations plan, which is a model decision (e.g., the demand of water at different wells for hydraulic stimulation depends on the development plan). That is why, in the most general case, the demand and/or production of water, oil, NGL (natural gas liquids), LNG (liquefied natural gas), methanol, liquid hydrogen, ammonia, or any other substance in liquid state, may also be model variables (Cafaro and Grossmann 2020). Their values depend on the operations plan to be determined by the model, while the pipeline network should be optimally sized together with that plan. Following industry practices, the pipeline diameter to select usually belongs to a set of available commercial diameters with specific costs per unit length (Bragalli et al. 2012; Araya et al. 2018), but there are works that relax the latter assumption (Cafaro and Grossmann 2014a).

Another common assumption in the design of liquid pipeline networks is the simplification of hydraulic and pressure calculations. Although they are relevant variables in daily operations, from the standpoint of a process designer they might be simplified. This is based on the incompressibility assumption for liquid flows at relatively low pressures, and on the cost of centrifuge pumps, which is generally low in comparison to the pipelines layout.

According to Cafaro and Grossmann (2020), liquid pipeline designers usually allow a maximum head loss per unit of length ($s_L = h_L/L$), thus leading to a maximum flow rate for every link i–j that increases with the pipeline diameter ($d_{i,j}$). In the following sections we present the details of this relationship for two particular cases of the energy industry: (1) the design of oil, NGL, and multiproduct pipelines, and (2) water pipeline networks for the development of unconventional resources.

Journal : SmallExtended 11750	Article No : 635	Pages : 30	MS Code : 635	Dispatch : 13-6-2022

3.1.1 Oil and refined product pipeline networks

When sizing liquid phase pipelines, petroleum engineers seek to avoid corrosion, erosion and water hammer effects, among other negative phenomena associated with the flow velocity. Based on that, they usually impose a maximum mean velocity (vmax_{LP}) of around 1.5 m/s (Society of Petroleum Engineers 2006). Equation (11) determines the maximum flow for oil and other liquids when linking separation facilities (i) with delivery points (j) through a pipeline of diameter $d_{i,j,t}$ installed in time period t. In other words, in the oil industry the maximum admissible flow is usually assumed to be directly proportional to the pipeline cross section.

$$TC^{\text{oil}}_{i,j,t} = \text{vmax}_{\text{LP}} \cdot \frac{\pi}{4} \cdot d^2_{i,j,t} \quad \forall i \in I, j \in J, t \in T. \tag{11}$$

In the particular case of ethane and other natural gas liquids (NGL), Cafaro and Grossmann (2014a) propose an equation to convert from volume to mass units. In essence, the fundamentals are the same since the liquid density is assumed to be constant regardless of the pipeline pressure.

However, in a more realistic problem representation, pipeline diameters need to be selected from a finite set of alternatives, as proposed by Drouven and Grossmann (2016). Then, Eq. (11) can be modified to impose a different transportation capacity according to the diameter of the pipeline installed between i and j in time period t. That is represented by Eq. (12), where $x_{i,j,d,t}$ is a 0–1 variable that equals 1 if and only if a pipeline with diameter δ_d is installed between i and j at period t.

$$TC^{\text{oil}}_{i,j,t} = \sum_d tc_{i,j,d} \cdot x_{i,j,d,t} = \sum_d \text{vmax}_{\text{LP}} \cdot \frac{\pi}{4} \cdot \delta^2_{i,j,d} \cdot x_{i,j,d,t} \quad \forall i \in I, j \in J, t \in T. \tag{12}$$

Assigning binary variables to identify individual alternatives for the pipelines, pumps, compressors and any other element in the network is a way of circumventing the difficulty posed by economies of scale and nonconvex cost functions. The fact that purely nonlinear programming models may not represent the pipeline network design adequately leads to the underlying MINLP model that is usually associated with these problems (Duran and Grossmann 1986). From now on, the equations to be presented assume that the pipeline diameter is a continuous decision variable, but adaptations to the discrete case (as in Eq. 12) are straightforward.

Regarding the calculation of capital expenditures, Cafaro and Grossmann (2014a) propose an economy of scale function with the pipeline diameter, which in their model is indeed a continuous decision variable, as presented in Eq. (13). Note that Capex is a non-negative variable determined by the pipeline diameter that is selected by the model.

$$\text{Capex}\left(d_{i,j,t}\right) = \alpha l_{i,j} d^{\beta}_{i,j,t} \quad \forall i \in I, j \in J. \tag{13}$$

In their MINLP formulation, they substitute $d^2_{i,j,t}$ by $dl_{i,j,t}$ in Eq. (11) to keep a linear form in all model constraints, while Eq. (13) is replaced by (13a) and directly introduced in the objective function.

$$\mathrm{Capex}\left(dl_{i,j,t}\right) = \alpha l_{i,j}\ dl_{i,j,t}^{\beta/2} \quad \forall i \in I, j \in J. \tag{13a}$$

Note that α and β are given parameters, with $0 < \beta < 1$, leading to a concave cost function. $l_{i,j}$ is the length of the pipeline segment connecting i and j, which is given, since the location of the nodes is known beforehand. Pump investment and installation costs are usually included in this function.

Finally, guidelines on how to calculate the operating costs (Opex) in every pipeline segment for a given pipeline diameter d can be found in Cafaro et al. (2015). The authors address the pumping cost for multiproduct liquid pipelines as a function of the flowrate. Their model assumes that batches move into the pipeline in plug flow, and interface or "transmix" volumes are neglected. From the relationship between friction losses and pumping rates by means of Darcy's law (Darcy, 1857) and Colebrook-White correlation, a nonlinear equation can be used to calculate the power PW that is required to compensate the friction loss according to the pipeline diameter and the flowrate at period t. Such a nonlinear equation is then integrated into the MINLP model (see "Appendix").

If the parameter ec_t represents the unit energy cost during the time period t, and Δt is the period length, pumping charges are given by Eq. (14), which also includes a linear term accounting for the elevation difference between nodes i and j. Note that $\Delta z_{i,j} = -\Delta z_{j,i}$, which implies that the cost of pumping fluids in one direction may be different than in the opposite direction, leading to different operating costs. Also, note that the latter term (due to gravity) is usually much smaller than the pressure drop due to friction in large pipeline networks.

$$\mathrm{Opex}\left(F_{i,j,t}\right) = ec_t\left[\mathrm{PW_L}\left(F_{i,j,t}, d\right) + F_{i,j,t}\rho g \Delta z_{i,j}\right]\Delta t \quad \forall i \in I, j \in J, t \in T. \tag{14}$$

3.1.2 Water pipeline networks

In the particular case of water pipeline networks, the Hazen–Williams correlation presented in constraint (15) is typically used for design purposes. It assumes that the maximum admissible flowrate follows a power function with the pipeline diameter.

$$TC_{i,j,t}^{\mathrm{water}} = k_{i,j} d_{i,j,t}^{\omega} \quad \forall i \in I, j \in J. \tag{15}$$

$k_{i,j}$ is a coefficient that depends on the rugosity of the internal walls of the pipeline and the head loss per unit length of pipe (s_L). Note that, in contrast to oil pipeline optimization problems, the unit head loss s_L is usually selected a priori, typically around 10 Pa/m (Caballero and Ravagnani 2019). The exponent ω is empirically determined by Hazen-Williams at the value of 2.6298.

Regarding energy consumption (operating costs), the selection of a fixed s_L for design purposes yields a simplified, linear formula. This implies that the flowrate is set at a value such that the head loss per unit length equals s_L. If the pipeline diameter is larger than required ($F_{i,j,t} < TC_{i,j,t}^{\mathrm{water}}$) then the pipeline will be idle for some time, but the flowrate during active intervals will be fixed at the value mentioned above. Hence, the pumping costs can be calculated from the friction head loss and elevation difference between the nodes ($\Delta z_{i,j}$), as in Eq. (16).

$$\mathrm{Opex}\left(F_{i,j,t}\right) = ec_t \rho g\left(s_L l_{i,j} + \Delta z_{i,j}\right)\eta F_{i,j,t}\Delta t \quad \forall i \in I, j \in J, t \in T. \tag{16}$$

It is interesting to note that by fixing the head loss per unit length s_L, the liquid transportation cost is a non-negative variable independent of the pipeline diameter that grows linearly with the flowrate.

3.2 Gas pipeline networks

The first MINLP model to address the optimal design of natural gas pipelines was proposed by Duran and Grossmann (1986). The authors solve an integrated problem in which both the configuration and sizing variables must be selected by the same model. In their seminal approach, the transportation capacity of a segment *i–j* with diameter $d_{i,j,t}$ is assumed to be given by the Weymouth (1912) correlation, as in Eq. (17).

$$TC^{gas}_{i,j,t} = \gamma^{-0.5} l_{i,j}^{-0.5}\left(P_{i,t}^2 - P_{j,t}^2\right)^{0.5} d_{i,j,t}^{2.667} \quad \forall i \in I, j \in J, t \in T, \tag{17}$$

where

$$\gamma = sg\theta\left[P_o/\left(0.375\,T_o\right)\right]^2. \tag{18}$$

s_g is the gas specific gravity at standard conditions and θ is the average gas temperature. Note that if the input and output pressures (P_i, P_j) are assumed to be known, the gas transportation capacity is a function of the pipeline diameter raised to 2.667. In further sections, we present a generalized model where this assumption is relaxed.

If the pipeline diameter is assumed to be a continuous variable, one may use an economy of scale function like (13) to determine the capital cost of the gas pipeline *i–j* installed at period *t*. In that case, by substituting $d_{i,j,t}$ with the variable $dg_{i,j,t}{=}d_{i,j,t}{}^{2.667}$, Eq. (17) is converted into a linear constraint, and the capital cost can be approximated by Eq. (19).

$$\mathrm{Capex}\left(dg_{i,j,t}\right) = \alpha l_{i,j}\, dg_{i,j,t}^{\beta/2.667} \quad \forall i \in I, j \in J, t \in T. \tag{19}$$

Drouven and Grossmann (2016) present an adapted form of Eq. (19) to the case in which the pipeline diameters are selected from a finite set of commercial sizes, as shown in Eq. (20).

$$\mathrm{Capex}\left(x_{i,j,d,t}\right) = \sum_d l_{i,j}\alpha\delta_d^{\beta} x_{i,j,d,t} \quad \forall i \in I, j \in J, t \in T. \tag{20}$$

Another interesting note on that work is that a simplified strategy to size gas pipelines is adopted, based on imposing an upper bound to the fluid velocity. The authors suggest that as a rule of thumb, operators strive to ensure that the fluid velocity in gas lines does not exceed 20 m/s to minimize noise emissions and to allow for corrosion inhibition. Thus, a preliminary pipeline sizing is made by relying on that simplified, maximum-velocity specification. However, it is also advisable to use standard gas flow equations such as the Weymouth or Panhandle correlations to calculate the pressure drops along individual pipeline segments after preliminary sizing. If these are beyond tolerable specifications, a larger diameter pipeline may be selected.

3.2.1 Fixing pressures at the nodes for design purposes

In many contributions addressing the optimal design of gas pipeline networks, usually based on a fixed number of echelons, reference values for inlet and outlet pressures at the segments are given beforehand to simplify the pipeline sizing equations (Cafaro and Grossmann 2014a; Montagna et al. 2021). The values in Fig. 4 are reference pressures proposed by Montagna et al. (2021) for the optimal design of unconventional oil and gas gathering networks. The underlying assumption is that the pressure at the wellheads is high enough to make oil, gas and water flow towards the tank batteries for separation. After the combined flow is separated, the gas flow is pressurized at the compressors to supply the delivery points in compliance with the requirements. It is important to note that all intermediate pressures are arbitrarily assumed at some reference level for design purposes (see Fig. 4). However, the actual pressure drops need to be finally determined after the optimization is performed, considering the actual flowrates between nodes. Computation of actual pressures is usually made by detailed flow simulations (Chuen et al. 2017). In the most general case, only wellpads and compressors (inlet/outlet) pressures are given, while the others stand for optimization variables that may help to achieve better solutions, as explained in the following section.

3.2.2 Optimizing pressures to improve pipeline utilization

In a more general problem, an additional challenge is to track pressures at the nodes of the gas pipeline network over time. As expressed by Eq. (17), the pipeline transportation capacity for compressible fluids can be modified by handling pressures. If the number of segments along which the flow of gas is directed to reach a destination node depends on the network design (i.e., the number of echelons is not fixed beforehand) one should accurately define the inlet and outlet pressures at every segment for every time period. By converting gas pressures into decisions variables to be determined, gas

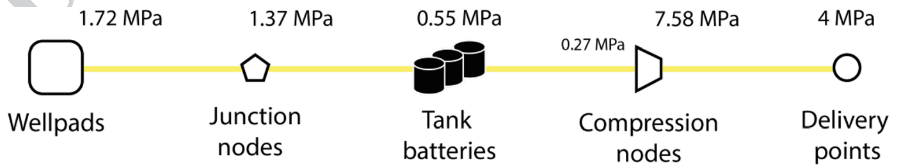

Fig. 4 Reference pressures in 10^6 Pascal (MPa) at different nodes of a network gathering unconventional oil and gas (Montagna et al. 2021)

flowrates and flow directions can be optimally handled along the time horizon to make a better use of the pipeline transportation capacity. Needless to say, though, the mathematical formulation becomes more complex.

If we raise both sides of the Weymouth correlation presented in Eq. (17), we derive a quadratic form as in Eq. (21). In this equation, $d_{i,j,t}$ is the diameter of the pipeline connecting i to j, already installed at period t.

$$\left(TC^{\text{gas}}_{i,j,t}\right)^2 = \gamma^{-1} l^{-1}_{i,j} d^{5.334}_{i,j,t}\left(P^2_{i,t} - P^2_{j,t}\right) \quad \forall i \in I, j \in J, t \in T. \tag{21}$$

To simplify the formulation, we introduce the variable $P^{\text{sq}}_{i,t} = P^2_{i,t}$ to account for the square pressure at the node i during period t, which is the actual variable to be optimally set by the model. Note that because pipeline flows can be reversed, it is necessary to enforce the transportation capacity to be zero when the difference of square pressures is negative (the gas flows in the opposite direction). This is imposed by Eqs. (22)–(24), where $u_{i,j,t}$ is a binary variable that takes value one if the gas flows from i to j during time period t, and zero otherwise.

$$u_{i,j,t} + u_{j,i,t} = 1 \quad \forall i \in I, j \in J, t \in T, \tag{22}$$

$$\Delta P^{\text{sq}}_{i,j,t} \le \left(P^{\text{sq}}_{i,t} - P^{\text{sq}}_{j,t}\right) + \Delta sp^{\text{Max}}_{i,j}\left(1 - u_{i,j,t}\right) \quad \forall i \in I, j \in J, t \in T, \tag{23}$$

$$\Delta P^{\text{sq}}_{i,j,t} \le \Delta sp^{\text{Max}}_{i,j} u_{i,j,t} \quad \forall i \in I, j \in J, t \in T. \tag{24}$$

$\Delta \text{sp}^{\text{Max}}_{i,j}$ is the maximum difference of square pressures for the pipeline segment i–j, usually given by the difference between the maximum square pressure at the source i and the minimum square pressure admitted at the inlet of a junction, compressor or processing facility.

Given that the gas flows, by definition, are greater or equal to 0, the maximum transportation capacity of a pipeline at a certain time period t can be imposed through Eq. (25).

$$\left(F_{i,j,t}\right)^2 \le \gamma^{-1} l^{-1}_{i,j} d^{5.334}_{i,j,t} \Delta P^{\text{sq}}_{i,j,t} \quad \forall i \in I, j \in J, t \in T. \tag{25}$$

Interestingly enough, if the pipeline diameter is selected from a finite set of alternatives, the nonlinear inequality (25) yields a convex quadratic constraint. Then, the mathematical formulation for the gas pipeline network optimization results in a mixed-integer quadratically constrained programming (MIQCP) model with a convex relaxation. Modern solvers like Gurobi 9.0 (Gurobi Optimization LLC 2021) are able to solve moderate size instances of this problem to global optimality in reasonable CPU times (Montagna et al. 2022).

3.3 Multiphase pipeline networks

One of the most difficult challenges in the optimal design of pipeline networks is when they carry multiphase flows, combining liquid and gas phases. That is the case of "flowlines" in the oil and gas industry, which gather the wells production (comprising oil, gas and water) to separation facilities. Even though there may be more than two phases in the liquid stream (e.g., oil and water, which are immiscible), they are usually simplified into a single liquid phase for pipeline sizing purposes (Montagna et al. 2021). Similarly, for simplicity, in this work we only consider two aggregate phases: liquid and gas.

When designing multiphase pipeline networks, the difficulty to calculate pressure drops comes from the diversity of flow patterns. Figure 5 shows the flow patterns that are expected according to the velocities of the gas and liquid streams. That is why pressure drops are usually calculated from empirical correlations. One of the most widely used procedures is due to Lockhart and Martinelli (1949), while the Society of Petroleum Engineers (2006) suggests simplified guidelines for multiphase pipeline sizing that are particularly useful for shorter pipeline segments. Both procedures are detailed in the following section and in the "Appendix".

3.3.1 Multiphase pressure drop prediction and pipeline sizing guidelines

The Lockhart–Martinelli procedure (LM) is a widely used method to compute pressure drops on multiphase (liquid and gas) pipelines based on empirical correlations (Lockhart and Martinelli, 1949). The approach aims to obtain the overall pressure drop in a straight pipeline segment from the expected pressure drops of the individual phases, i.e., as if they were flowing alone. To achieve this, the method computes an intermediate parameter (X_{LM}) as the square root of the ratio between isolated pressure drops, which is subsequently entered on a particular function (Wilkes 2005). This function fits empirical data according to the flow regime of the multiphase fluid composition, yielding the coefficient Y_G. Finally, the coefficient Y_G is multiplied by the gas pressure drop to obtain the overall, multiphase pressure drop.

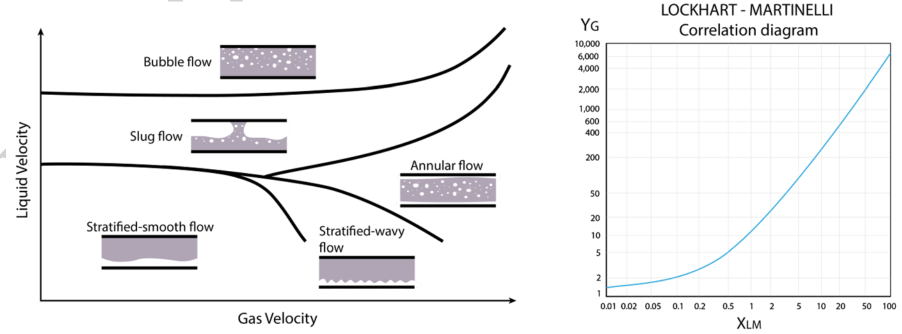

Fig. 5 Two-phase flow patterns and Lockhart–Martinelli correlation diagram to calculate the pressure drops in liquid–gas pipelines

Figure 5 shows the correlation diagram that relates the LM parameter (X_{LM}) with the coefficient Y_{G} under turbulent flow regime. For more details, we refer the reader to the "Appendix".

On the other hand, the Society of Petroleum Engineers (2006) provides guidelines for the sizing of multiphase pipelines based on admissible flow velocities to prevent erosion, corrosion, noise and water hammer effects. The method imposes an upper bound to the liquid flow rate depending on the pipeline diameter and the gas to liquid ratio, among other physical properties. More details on the SPE guidelines are also given in the "Appendix".

4 Solution strategies

The nonlinear equations governing pressure drops, energy consumption and capital investments like the ones presented in this paper have motivated purely nonlinear (NLP) approaches and solution techniques, as reviewed by Mah and Shacham (1978). Several years later, the design of gas pipeline networks inspired one of the most valuable contributions to the solution of MINLP problems, namely the outer-approximation algorithm (Duran and Grossmann 1986). The inclusion of binary variables permitted to circumvent the difficulties posed by economies of scale, usually represented by concave cost functions. Such concave cost functions, particularly challenging when sizing pipelines, inspired the development of alternative functions to avoid unbounded gradients in the NLP relaxation (Cafaro and Grossmann 2014b). The latter idea was implemented in the seminal work on the optimal design of shale gas supply chains (Cafaro and Grossmann 2014a), where a Branch-Refine-Optimize (BRO) algorithm allowed to address large-scale instances of the problem.

In a more recent contribution, Cafaro and Grossmann (2020) present an iterative algorithm searching for better integer solutions in shorter computational times. The procedure implies separating the operations planning from the network design problem to obtain a first approximation to the optimum. If the production planning is prioritized, the required pipeline networks are rather expensive, with oversized capacities. Based on that fact, an interesting way to improve the solution in a second step is to solve the integrated problem but just considering the subset of connections suggested by the original, myopic strategy. Although this subset is relatively large when compared with the optimal design, it excludes many of the connections in the superstructure of alternatives, thus reducing the model size.

Following a similar strategy, Montagna et al. (2021) develop a bi-level decomposition procedure that comprises a series of approximations derived from the overall MINLP model to design pipeline networks with multiphase flows. If solved in a specified sequence, the approximations are able to find efficient solutions to the original problem in modest computational times. The strategy consists in defining NLP formulations comprising all the fluid dynamic equations to estimate the transportation capacity of the pipelines, and an MILP formulation for the network design. For any potential multiphase pipeline in the network, an NLP model is solved whose objective is to maximize the admissible flowrate. Such NLP formulations can be regarded as systems of nonlinear equations that can be solved separately, before the network design optimization. Due to their small size, hundreds of these NLP problems can be solved to optimality in few minutes. After that, the MILP approximation relies on the results of the NLP subproblems, taking the resulting transportation capacities as inputs of the pipeline design problem.

5 Illustrative examples

In this section, we illustrate the solutions obtained with the formulations described in this work by means of representative case studies. Case study 1, originally presented by Cafaro and Grossmann (2014a), involves the optimal design of a supply chain network over a shale gas exploitation area, covering more than 10,000 km^2. Several potential sites for well drilling are considered, assuming different productivity profiles and hydrocarbon composition throughout the region. A network structure with a fixed number of echelons is developed. Case study 2 provides powerful insights on the design of water distribution networks, and confirms the importance of coupling facility sizing decisions with those related to production planning. In contrast to the Case study 1, the network design problem does not impose a fixed number of echelons. A large case study involving 52 wells is described aiming to illustrate the potential of the mathematical formulation. Finally, Case study 3 presents the optimal design of the oil gathering network for unconventional production from a real-world field. The problem statement integrates several decisions such as tank batteries sizing and site selection, pipeline connections and diameters, location of junction nodes and multiperiod investments. A pre-processing NLP formulation is applied to estimate the maximum pipeline transportation capacity for multiphase flows across every possible segment in the superstructure, which is then included in the master MILP model.

5.1 Case study 1: Shale gas pipeline network

The first case study describes the optimal design of a supply chain with a fixed number of echelons in a real world shale play. In their work, Cafaro and Grossmann (2014a) seek not only to define the optimal design of a pipeline network divided in three echelons, but also to determine the optimal drilling and fracturing strategy that maximizes the NPV (Net Present Value) of the shale gas project, over a 40-quarter planning horizon. A base superstructure comprising a total of 29 nodes with known location is postulated, including 9 potential sites for developing wells, 8 potential sites for junction/compression of flows, 3 alternative sites for processing plants setup, 3 methane demanding nodes, 3 ethane demanding nodes, and 3 freshwater sources (see Fig. 6). All distances between potential nodes, required to compute pipeline lengths and water transportation paths for trucks, are measured in Euclidean norm.

The model accounts for several types of pipelines to transport three different fluids: shale gas, ethane and methane. It also considers the seasonality on natural gas prices and freshwater availability, as well as the economies of scale

Journal : SmallExtended 11750 | Article No : 635 | Pages : 30 | MS Code : 635 | Dispatch : 13-6-2022

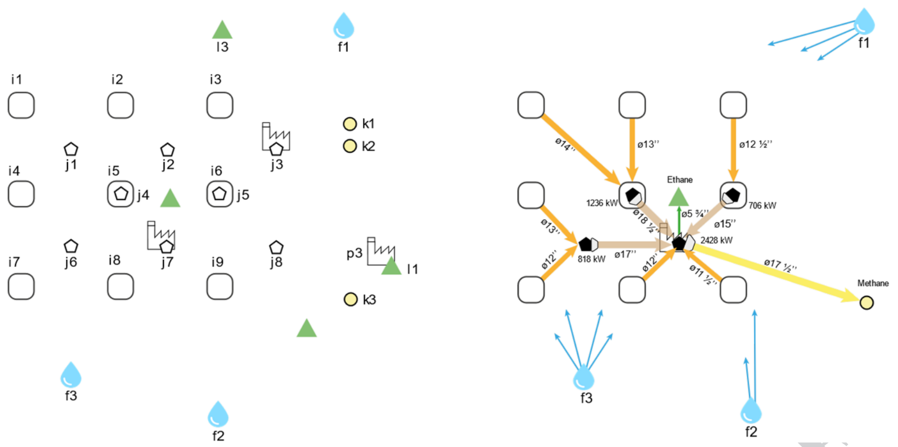

Fig. 6 Base superstructure of nodes and optimal pipeline network design for Case study 1, presented by Cafaro and Grossmann (2014a). Compressors power is given in kilowatts (kW)

representing the costs of pipelines, processing plants, wells and compressors. On the other hand, shale gas productivity at each well is approximated by a discrete-time decreasing power function of the well age. A variant of the same problem, including gas “wetness” variability (i.e., different amounts of ethane and NGL in the gas composition) according to the wellpad location is also addressed.

Due to the large size and the computational challenges of the MINLP formulation developed by the authors, a two-level solution strategy is proposed to efficiently tackle the problem. The Branch-Refine-Optimize (BRO) algorithm successively refines piecewise-linear underestimations of the concave cost functions. Since the original MINLP model seeks to maximize the NPV, such MILP relaxations are solved to find increasingly tighter upper bounds of the objective function. Feasible solutions are then obtained by solving a reduced MINLP model that omits network structure variables (potential connections) that take value zero in the optimal solution of the corresponding MILP approximation. The reduced MINLP then seeks to improve the lower bounds of the original model. The MILP-MINLP procedure is repeated until the lower bound (from the MINLP) and upper bound (from the MILP) are close enough to satisfy an optimization tolerance. Using the BRO algorithm, this case study is solved after eight major iterations (using GAMS software and GUROBI 5.5 as the MILP solver, and DICOPT 24.1, combining GUROBI 5.5 and CONOPT 3.15, as the MINLP solver). After 13.7 h of computation on an Intel® Core i7 CPU, 2.93 GHz, 12 GB RAM (with six parallel threads) the algorithm reports a global optimality gap of 2.5%. The first MILP model, with piecewise-linear approximations dividing the domain of concave cost functions into two sectors, is rather large, including 51,888 equations, 47,643 continuous variables and 3490 binary variables. It takes almost 5 h of CPU to reach an optimality gap of 0.25% in the same computer. The first MILP approximation is usually the one that takes the longest time within the BRO algorithm, since there is no integer-feasible solution to start from.

The optimal solution yielded by the MINLP formulation of Cafaro and Grossmann (2014a) is presented at the right of Fig. 6. Results suggest that only one shale gas processing plant is installed in the network. Moreover, optimal shale gas compression nodes are established at three junction sites. Finally, the processing plant is linked to a delivery point to supply methane at the required pressure.

The optimal solution also suggests that economies of scale are critical, leading to the installation of facilities and pipelines during the first period of the planning horizon, and not expecting any further expansion. Utilization of the processing and transportation facilities peak at time period 7, and remain high for many subsequent periods, as more wells are conveniently developed to steadily deliver the products. Indeed, important conclusions are drawn from this work on how to optimize reservoir development strategies in coordination with pipeline network designs. By doing so, high utilization levels and stable flows can be obtained over the network..

Finally, sensitivity analyses prove that the proposed design is certainly efficient, showing significant worsening of the objective function when some changes are enforced in the network configuration (e.g., adding a second processing plant). Furthermore, by extending the model capabilities to consider different gas compositions for different locations over the field, it is demonstrated that the optimal network design remains the same, providing important hints on the robustness of the solution.

5.2 Case study 2: Water distribution network for shale gas production

This example addresses the optimal planning of drilling, fracturing and completion operations together with the optimal design of pipeline networks supplying, processing and recycling water for hydraulic fracturing in unconventional gas plays. This is an illustrative case comprising 52 wells to be developed over 12 wellpads, with 2 potential sources of freshwater (see Fig. 7). The one-year planning horizon is discretized into 52 weekly periods, and there are alternative pipeline diameters and water storage tanks to install in the network. The pipeline connections among the nodes do not need to follow a predefined number of echelons, and flows in pipeline segments (links) may be reversed at different time periods. Due to the combinatorial complexity of the overall problem, Cafaro and Grossmann (2020) propose a two-stage decomposition algorithm. The first stage seeks to establish the optimal configuration of the water distribution network having obtained, in advance, the well development plan that maximizes the benefits from gas production. In other words, the water distribution network is subordinated to a so-called “myopic” production plan, from which one determines the water requirements for each of the wellpads over the time horizon. In the second stage, the full optimization model (simultaneously optimizing the production strategy and the network design) is solved, limiting the superstructure of alternatives to the ones suggested by the first stage.

The proposed MILP model for Case study 2 comprises 13,567 constraints, 10,586 continuous variables and 3251 binary variables. The computational time required for the first stage (myopic solution) is around 145 s, while the simultaneous optimization of well development and water network design and operation, based on the

Journal : SmallExtended 11750	Article No : 635	Pages : 30	MS Code : 635	Dispatch : 13-6-2022

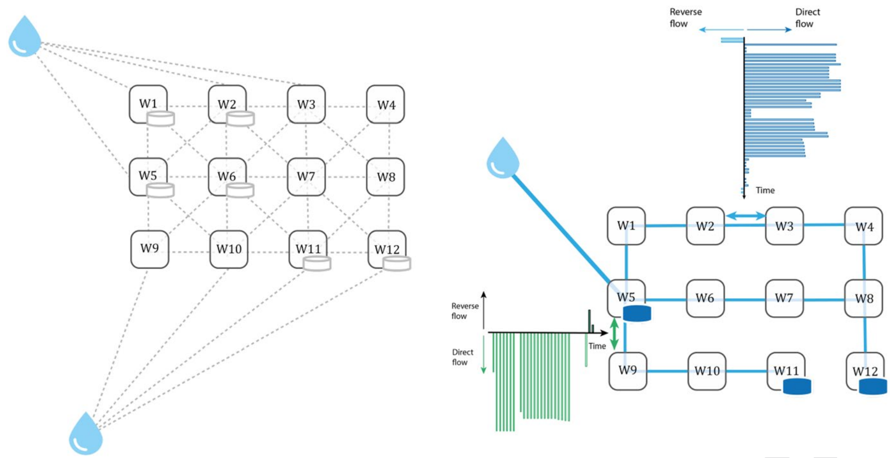

Fig. 7 Superstructure of alternatives for the water distribution network (left) and the optimal network design found by the model proposed by Cafaro and Grossmann (2020) that maximizes the Net Present Value of the project (right)

reduced superstructure yielded by the first stage, takes less than 10 min to converge. Instead, the full MILP model with all possible connections takes 27 h to close the optimality gap to 0.4% on an Intel® Core i7 CPU, 12 GB RAM. The best solution found is shown at the right of Fig. 7, and has been obtained using GAMS software with CPLEX 12.9 as the MILP solver, running on 4 parallel threads.

Results show that decoupling the operations planning from the pipeline network design may lead to very complex configurations, with numerous pipelines and storage facilities to satisfy water demand under myopic, intensive well fracturing plans. In addition, several flow reversals are necessary along the time horizon, leading to higher storage levels and operating costs. In contrast to myopic solutions, simultaneous optimization yields an improvement of 6% in the NPV (from 101.28 to 107.23 million dollars), mainly due to significant savings in water costs at the expense of a less intensive production plan. In fact, a substantial reduction in the water pipeline network complexity is observed, shortening the length of the pipelines by 27%, and reducing water storage capacities by 50%.

Finally, the robustness of the pipeline network is tested by admitting more flexible field development strategies, where the earliest times to start production in different wellpads are shortened. Despite the significant changes in the development plan, the structure of the water distribution network remains unchanged.

5.3 Case study 3: Shale oil gathering network

Montagna et al. (2021) address a real-world case study from the O&G industry involving the optimal design of pipeline networks to gather unconventional oil production. The authors assume that there is a pre-determined development plan from which operators predict the flows of oil/gas/water to be produced from each wellpad during the following 5 years. The planning horizon is divided into 60 monthly

Journal : SmallExtended 11750	Article No : 635	Pages : 30	MS Code : 635	Dispatch : 13-6-2022

periods and gas-to-oil, and water-to-oil ratios are assumed to remain constant in time and independent of the wellpad location. In addition, every wellpad is specified as a potential location for pipeline junction nodes, while potential tank battery locations are also proposed as possible junction points from where the production can be transported to the Centralized Delivery Points (CDP). The base superstructure restricts pipeline connections to a total of four echelons: wellpads–junctions, junctions–tank batteries, tank batteries–junctions, and junctions–CDP (see Fig. 8). The first two segments carry multiphase flows, while the last two transport single phases, namely oil, gas and water in separate pipelines.

To determine the transportation capacity of multiphase pipelines according to the length, diameter and flow composition, roughly 2000 NLPs (each involving 21 variables and 15 constraints) are solved using GAMS/IPOPT 30.3 for every segment

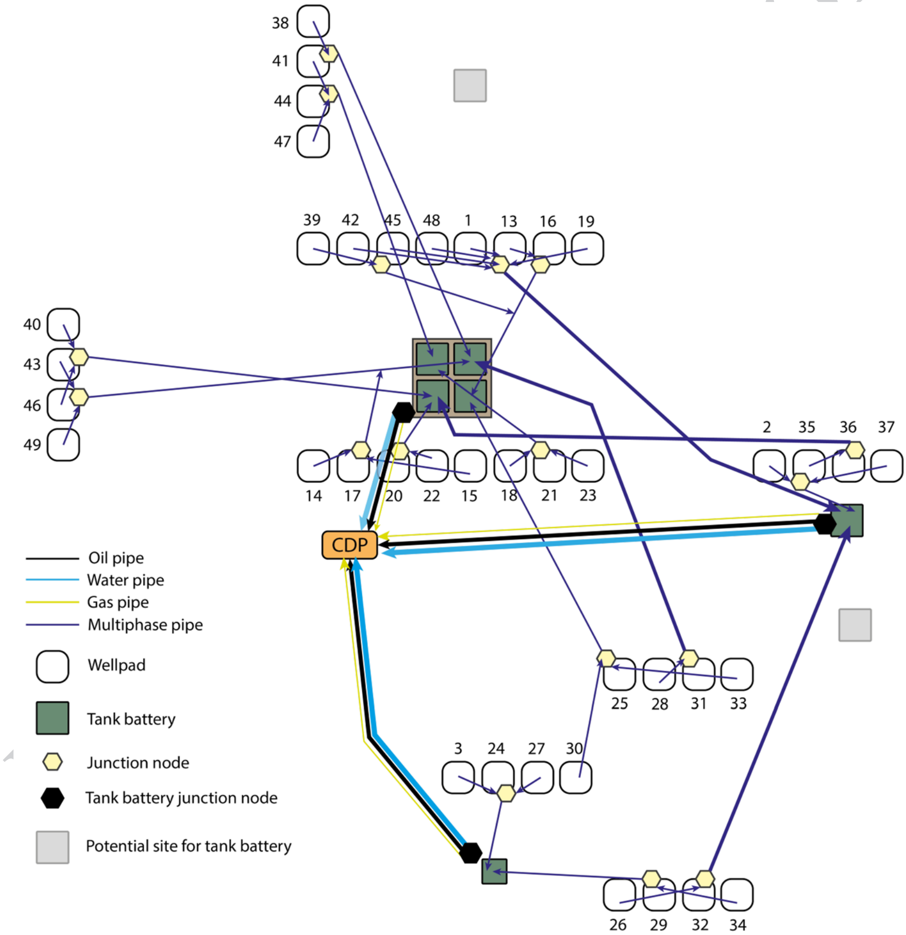

Fig. 8 Gathering network design obtained by Montagna et al. (2021) using a tailored solution algorithm. The number over each wellpad represents the start period of its production, in months

in the superstructure. As described in Sect. 4, the NLP models are based on the Lockhart–Martinelli correlation and SPE guidelines, assuming fixed pressures at the intermediate nodes of the network. After solving this pre-processing stage in about 15 min, conservative maximum flowrates are imposed for every triplet origin–destination-diameter, yielding a MILP approximation of the original MINLP that involves 2250 discrete variables, 24,600 continuous variables and 29,441 constraints after pre-processing. The authors propose an iterative solution algorithm aiming to sequentially provide tighter bounds for the number of tank batteries to be installed and enhance convergence. Making use of the algorithm and after 10 h of computation using the software GAMS/GUROBI 9.0 as the MILP solver, a global optimality gap of 5.2% is achieved, and the best solution found is the one depicted in Fig. 8. All computations have been carried out on an Intel® Core i7 at 3.9 Ghz CPU, with 16 GB RAM, with no parallelization.

The gathering network comprises a total of 6 tank batteries, 17 junction nodes spread over the different areas and 3 battery junction points. Different pipeline diameters are selected across the network depending on the number of phases in the flow, the transportation distances, and the pressure requirements at the terminal nodes (see thickness of arrows in Fig. 8). Note that despite the apparent complexity of the network design, with many pipeline crossings, such connections are required to keep all the tank batteries with a sufficiently high utilization level, providing substantial savings on facility installation costs. For more details on the timing of investments along the planning horizon and other features of this solution, the reader is referred to the work by Montagna et al. (2021).

Although the solution obtained for this case study may still leave room for improvement (the optimality gap is slightly above 5%), it is the first formulation published in the literature that shows the great potential to size pipelines and manage multiphase flows in the design of pipeline networks. More importantly, it allows obtaining near optimal solutions in reasonable times. The authors conclude that the simultaneous design of the oil gathering network and the definition of production strategies could bring substantial savings to the total investment costs, also stabilizing resource utilization.

A summary of the model types and dimensions, solution strategies and computational results for the three case studies is given in Table 1.

6 Concluding remarks

We have presented a systematic classification of the mathematical models proposed in recent years to the optimal design and operation of pipeline networks in the energy industry. To address this problem, the first important question to answer is how to connect the elements of the network, selecting the links from a superstructure of alternatives. We distinguish among two different approaches from the topological perspective: networks with a fixed number of echelons, and networks where the number of echelons is not determined a priori. The second model structure is more flexible, making it possible to reverse the flow direction in some of the segments/links at different time periods to make a better use of the pipelines. In both

Journal : SmallExtended 11750 | Article No : 635 | Pages : 30 | MS Code : 635 | Dispatch : 13-6-2022

Table 1 Problem features, model types, dimensions, solution strategies and computational results for the illustrative case studies

Case Study	Network Design Problem	Fluids	Superstructure size		Solution Strategy	Instance	Model Type	Model Size			Solver	CPU time	Global optimality gap	Hardware
			Nodes	Echelons				Eqs	Cont. Vars	Int Vars				
1	Shale gas supply chain	Shale gas, ethane, methane	29	3	Branch-Refine-Optimize (BRO) Algorithm	Full problem	MINLP	45,628	44,513	360	BARON 13	24 h	No Solution	Intel Core i7 CPU 2.93 Ghz, 12 GB RAM, 6 parallel threads
					Iterative: MILP + MINLP	Piece-wise linear approximation	MILP	51,888	47,643	3,490	GUROBI 5.5/ ICOPT 24.1 with GUROBI 5.5 and CONOPT 3.15	13.7 h	2.50%	
2	Water distribution network for shale gas production	Water	14	Undetermined	Decoupling Development Planning and Network Design Iterative: Reduced MILPs	Field Development Strategy, then Network design	MILP	13,567	10,586	1,889	CPLEX 12.9	145 s	0%	Intel Core i7 CPU 2.5 Ghz, 16 GB RAM, 4 parallel threads
						Simultaneous Network design and field development	MILP	13,567	10,586	3,251	CPLEX 12.9	27 h	0.40%	

Journal : SmallExtended 11750 | Article No : 635 | Pages : 30 | MS Code : 635 | Dispatch : 13-6-2022

Table 1 (continued)

Case Study	Network Design Problem	Fluids	Superstructure size		Solution Strategy	Instance	Model Type	Model Size			Solver	CPU time	Global optimality gap	Hardware
			Nodes	Echelons				Eqs	Cont. Vars	Int Vars				
3	Shale oil gathering network	Multiphase, oil, water natural gas	46	4	Decomposition of MINLP into NLP (maximum admissible flow) and MILP (network design)	Full problem	MINLP	59,441	66,600	2,250	BARON 20	24 h	No Solution	Intel Core i7 CPU 3.9 Ghz, 16 GB RAM, no parallel mode
					Two-Stage: NLP + MILP	Maximum admissible flows	NLP	15	21	–	IPOPT 30.3	< 1 s	0%	
						Pipeline network design	MILP	29,441	24,600	2,250	GUROBI 9.0	10 h	5.20%	

cases, the location of the existing and potential nodes is given and the topology of the network is solved by means of linear integer programming models, in which the main decisions are represented by 0–1 variables accounting for the installation of a pipeline segment between a pair of nodes. The second question is about the size (diameter) of the pipeline to install between the nodes. Sizing pipelines implies the use of fluid dynamics equations to predict the pressure drops, which are strongly nonlinear and dependent on the fluid state. Moreover, economies of scale functions are integrated into these models, yielding mixed-integer nonlinear formulations that are computationally challenging.

We have reviewed recent contributions that can be roughly classified according to the fluid to transport in three groups: liquid pipeline networks, gas pipeline networks, and multiphase pipeline networks. For each of these categories, we have presented alternative models and equations to estimate the pipelines transportation capacity as well as their capital and operating costs. We have described relevant applications of liquid pipeline networks in the energy industry like oil gathering networks, oil products transmission pipelines and water distribution networks for well stimulation. In turn, we have analyzed recent contributions to the optimal design and operation of natural gas pipeline networks, showing that they can be properly represented by mixed-integer quadratically constrained programming (MIQCP) formulations, which is a promising feature given the significant advances of the solution algorithms devised for this particular kind of nonlinear problems in the last few years. Finally, we have highlighted current trends in the optimal design and operation of multiphase pipeline networks, the most complex type of problems that are faced, for instance, when sizing the flowlines that carry a mixture of oil, gas and water from the wellbores to separation facilities over oil and gas fields.

The optimal design of pipeline networks has inspired process systems engineers and operations researchers since the early days of mathematical programming. It is not surprising to see that even today it is an active research field that continues inspiring new modeling frameworks and solution strategies. The transportation of liquid and/or gas hydrogen, ammonia, bioethanol, biogas (or renewable natural gas) and any other energy carrier produced from renewable sources represents a new frontier, which will certainly benefit from the developments described in this work.

Appendix

Calculation of energy consumption in a pipeline segment

A method to compute pumping cost for multiproduct pipelines as a function of the flowrate is presented by Cafaro et al. (2015), under the assumption of batch flow and neglecting "transmix" volumes. The mean velocity (U) in a straight segment is given by the ratio between the pump rate and the pipeline section, while the Reynolds number is given by $\mathrm{Re} = 4F/\pi d\nu$, with ν being the fluid kinematic viscosity. If refined products flow in turbulent regime into the pipeline segments ($\mathrm{Re} > 4 \times 10^5$) the relationship between the head loss due to friction (h_L) and the pump rate can be derived from the Darcy's equation (Darcy 1857) in Eq. (26). This equation introduces the dimensionless friction factor f. Indices of nodes and time periods are omitted for simplicity.

$$h_L = f\frac{l}{d}\frac{U^2}{2g} = 8f\frac{l}{d^5}\frac{F^2}{g\pi^2}. \tag{26}$$

Moreover, the friction factor f can be calculated through the Colebrook-White equation, as in Eq. (27).

$$\frac{1}{\sqrt{f}} = -2\log_{10}\left(\frac{\varepsilon/d}{3.7} + \frac{2.51}{\mathrm{Re}\sqrt{f}}\right). \tag{27}$$

Equation (27) involves an implicit function accounting for two contributions: the pipeline wall roughness (ε), and the flow turbulence. Finally, the power required to compensate the friction loss is given by Eq. (28), representing a nonlinear function rapidly increasing with the flowrate F.

$$\mathrm{PW}_L = \frac{h_L F \rho g}{\eta} = \frac{8\rho}{\pi^2\eta}\frac{l}{d^5} f\, F^3. \tag{28}$$

The parameter ρ is the liquid density, g is the gravitational constant, l is the length of the pipeline and η is the pump efficiency. Equations (29) and (30) are introduced in the MINLP model presented in Sect. 3.1.1 to calculate the energy consumption in every pipeline segment over time period t.

Lockhart–Martinelli procedure to compute pressure drops in multiphase flows

The Lockhart–Martinelli (LM) procedure has been devised to predict the pressure drop for fully developed gas–liquid flows. The first step is to calculate the individual effects of liquid and gas phases, i.e., the pressure drops ΔPl and ΔPg that would be expected if the liquid and gas streams were flowing alone through the same pipeline. The second step is to obtain the LM parameter (X_{LM}), as in Eq. (29).

$$X_{\mathrm{LM}} = \sqrt{\frac{\Delta Pl}{\Delta Pg}}. \tag{29}$$

Finally, the overall pressure drop in the pipeline segment is estimated from the pressure drop of the gas phase, as in Eq. (30).

$$P_{i,t} - P_{j,t} = Y_G \Delta Pg = \left(1 + X_{\mathrm{LM}}^{2/n}\right)^n \Delta Pg. \tag{30}$$

Equation (30) follows the Wilkes function (Wilkes 2005) to fit the data of the LM empirical results. The parameter n depends on the flow regime of each phase, and is equal to 4.12 if both liquid and gas phases flow in turbulent regime, as usually seen in industrial applications. In horizontal pipelines, the pressure drop of the liquid phase ΔPl can be obtained from the Darcy equation (see Eqs. (31) and (32)), while the gas pressure drop ΔPg is related to the pipeline diameter and the gas flow through the Weymouth correlation (see Eq. (17) in Sect. 3.2). Note that the flow rates of the two different phases need to be tracked through separate variables ($F_{\mathrm{LIQ},i,j,t}$ and $F_{\mathrm{GAS},i,j,t}$), as explained in Sect. 2.1.

SPE guidelines for sizing multiphase pipelines

Similar to the liquid pipeline sizing problem, the SPE (2006) suggests maximum velocities for multiphase flows to prevent erosion, corrosion, noise or water hammer effects. Reference maximum velocities are 18 m/s to inhibit noise and 15 m/s to prevent corrosion. A more accurate procedure to obtain maximum velocities is suggested as follows.

1. Obtain the average multiphase density (ρ_{avg}) from the gas to liquid ratio (glr) as:

$$\rho_{\mathrm{avg}} = \frac{12409\ \mathrm{sl}\ P_{\mathrm{in}} + 2.7\ \mathrm{glr}\ \mathrm{sg}\ P_{\mathrm{in}}}{198.7 P_{\mathrm{in}} + z\ \mathrm{glr}\ \theta}, \tag{31}$$

   where z is the gas compressibility factor, r the gas/liquid ratio [ft^3/bbl], θ the flow temperature [°R], P_{in} the inlet pressure in PSI, *sl* the specific gravity of the liquid phase relative to water, and sg the specific gravity of the gas, relative to air.
2. Set the maximum velocity for the liquid phase as:

$$v_{\max} = \rho_{\mathrm{avg}}^{-0.5}, \tag{32}$$

   where ζ is an SPE specific constant with a value of 150 for solids-free fluids and continuous service operation, and $v_{\max}$ is measured in ft/s.
3. Impose an upper bound on the liquid flowrate, as shown in Eq. (33).

$$F_{\mathrm{LIQ},i,j,t} \leq \xi v_{\max} d_{i,j,t}^2, \tag{33}$$

   with $\xi = 0.64516\left(11.9 + z\ \mathrm{glr}\ \theta / 16.7 P_{\mathrm{in}}\right)^{-1}$.

Acknowledgements Financial support from National University of Litoral under Grant CAI+D 2020 50620190100163LI, and CONICET under Grant PIP 11220200103053CO is gratefully acknowledged. Financial support from the Center of Advanced Process Decision-making at Carnegie Mellon University is also gratefully acknowledged.

References

Araya N, Luca FA, Cisternas LA, Gálvez ED (2018) Design of desalinated water distribution networks: complex topography, energy production, and parallel pipelines. Ind Eng Chem Res 57:9879–9888

Bhaskaran S, Salzborn FJM (1979) Optimal design of gas pipeline networks. J Oper Res Soc 30:1047–1060

Bragalli C, D'Ambrosio C, Lee J, Lodi A, Toth P (2012) On the optimal design of water distribution networks: a practical MINLP approach. Opt Eng 13:219–246

Caballero JA, Ravagnani MASS (2019) Water distribution networks optimization considering unknown flow directions and pipe diameters. Comput Chem Eng 127:41–48

Cafaro DC, Cerdá J (2012) Rigorous scheduling of mesh-structure refined petroleum pipeline networks. Comput Chem Eng 38:185–203

Cafaro DC, Grossmann IE (2014a) Strategic planning, design, and development of the shale gas supply chain network. AIChE J 60:2122–2142. https://doi.org/10.1002/aic.14405

Cafaro DC, Grossmann IE (2014b) Alternate approximation of concave cost functions for process design and supply chain optimization problems. Comput Chem Eng 60:376–380. https://doi.org/10.1016/j.compchemeng.2013.10.001

Cafaro DC, Grossmann IE (2020) Optimal design of water pipeline networks for the development of shale gas resources. AIChE J. https://doi.org/10.1002/aic.17058

Cafaro VG, Cafaro DC, Mendez CA, Cerdá J (2015) MINLP model for the detailed scheduling of refined products pipelines with flow rate dependent pumping costs. Comput Chem Eng 72:210–221. https://doi.org/10.1016/j.compchemeng.2014.05.012

Chuen CM, Nukman MA, Mandal D, Salim AS, Sepulveda W, Vaca JC, Goh K (2017) Application of allocation algorithm for surveillance and optimization of intelligent wells. In: SPE/IATMI Asia Pacific Oil & Gas Conference and Exhibition. https://doi.org/10.2118/186421-MS

Darcy H (1857) RecherchesExpérimentales Relatives au Mouve-ment de l'Eau dans les Tuyaux [Experimental Research on the Movement of Water in Pipes]. Mallet-Bachelier, Paris

Drouven M, Grossmann IE (2016) Multi-period planning, design, and strategic models for long-term, quality-sensitive shale gas development. AIChE J 62:2296–2323. https://doi.org/10.1002/aic.15174

Duran MA, Grossmann IE (1986) A mixed-integer nonlinear programming algorithm for process systems synthesis. AIChE J 32:592–606. https://doi.org/10.1002/aic.690320408

Gurobi Optimization LLC (2021) Gurobi optimizer reference manual

Hellemo L, Tomasgard A (2015) A generalized global optimization formulation of the pooling problem with processing facilities and composite quality constraints. TOP 24:409–444

Lockhart R, Martinelli RC (1949) Proposed correlation of data for isothermal two-Phase, two-component flow in pipes. Chem Eng Prog 45:38–48

Mah RSH, Shacham M (1978) Pipeline network design and synthesis. Adv Chem Eng 10:125–209

Montagna AF, Cafaro DC, Grossmann IE, Burch D, Shao Y, Wu X-H, Furman K (2021) Pipeline network design for gathering unconventional oil and gas production using mathematical optimization. Opt Eng. https://doi.org/10.1007/s11081-021-09695-z

Montagna AF, Cafaro DC, Grossmann IE, Ozen O, Shao Y, Zhang T, Guo Y, Wu X-H, Furman K (2022) Surface facility optimization for combined shale oil and gas development strategies. Comput Chem Eng 2022:2

Presser DJ, Cafaro VG, Cafaro DC (2022) Optimal sourcing, supply and development of carbon dioxide networks for enhanced oil recovery in CCUS systems. Comput-Aided Chem Eng 2022:5

Saldanha-da-Gama F (2018) Comments on: extensive facility location problems on networks: an updated review. TOP 26:229–232

Selot A, Kuok LK, Robinson M, Mason TL, Barton PI (2008) A short-term operational planning model for natural gas production systems. AIChE J 54:495–515

Society of Petroleum Engineers (2006) Petroleum engineering handbook AQ2

Weymouth TR (1912) Problems in natural gas engineering. Trans ASME 34:185–234

Wilkes JO (2005) Fluid mechanics for chemical engineers. Pearson, Upper Saddle River

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal:	**11750**
Article:	**635**

Author Query Form

Please ensure you fill out your response to the queries raised below and return this form along with your corrections

Dear Author

During the process of typesetting your article, the following queries have arisen. Please check your typeset proof carefully against the queries listed below and mark the necessary changes either directly on the proof/online grid or in the 'Author's response' area provided below

Query	Details Required	Author's Response
AQ1	Please provide MSC codes. For more details, if required, kindly visit http://www.ams.org/msc/.	
AQ2	Kindly provide necessary details for the ref. (Society of Petroleum Engineers 2006; Gurobi Optimization LLC 2021), if possible.	