Heliyon 10 (2024) e33947

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Review article

Optimization of petroleum products distribution via pipeline systems: Modeling and computational challenges

Rolando José Acosta-Amado

Departamento de Ingeniería Industrial, Facultad de Ingeniería, Diseño y Ciencias Aplicadas, Universidad Icesi, Cl 18 # 122-135, Cali 760031, Colombia

A R T I C L E   I N F O

Keywords:
Refined products distribution
Multi-product pipeline systems
Optimization

A B S T R A C T

Optimizing the distribution of refined oil products using pipeline systems has been posing a meaningful challenge to the operations research field of exceptional economic importance for the oil and gas industry. The solution of this problem stands on the interface of chemical engineering and operations research, the former has been the most important contributor, and the latter has been paying an increasing amount of attention to its solution over the last ten years or so. The goal of this work is *fourfold*: to unveil the current shape of the accomplished research work on this topical area according to its descriptive analytics, to present and discuss its modeling research perspectives, to outline its emerging research trends and to trigger discussions on its future research avenues. A future research agenda should study more realistic mathematical models for the system and solution procedures that exploit their algebraic structure.

1. Introduction

The United States have seen a dramatic increase in the oil production since 2007, which has come with an accelerated growth of its network of pipeline systems over the last 16 years or so. This kind of transportation mode includes oil, *petroleum products* and natural gas pipeline systems. According to the *American Petroleum Institute* (API) and the *Association of Oil Pipelines* (AOPL), by year 2013 this network spanned 192396 miles. Five years later, the total liquid pipeline infrastructure had grown over 23000 miles up to 215736 miles; a significant 12% increase with respect to year 2013. This increment in the size of the transportation infrastructure allowed a dramatic 44% increment in the total crude oil and *petroleum products* barrels delivered by pipelines [1]. Furthermore, the ever-growing trend on the mileage of pipelines infrastructure in the United States continued since year 2017 until year 2020 reaching some 229454 miles in length. Nonetheless, recent data suggests that this trend might have slackened afterwards (Fig. 1, [2]).

A steadily meaningful portion of this kind of infrastructure is composed by the petroleum products, or *multi product*, pipeline systems; often also called *poly-ducts*. Poly-ducts are a type of pipelines used for the transportation of refined products (e.g.: gasoline, diesel, jet fuel, kerosene, etc.) to consumer markets. They have proven to be both an efficient and safe alternative to move large amounts of this kind of freight across long distances. In year 2013 there were 65352 miles of poly-ducts that were used to deliver 6.642 billion barrels of petroleum products (gasoline, jet fuel, diesel, etc.) and natural gas liquids (propane, ethane, butane, etc.). It was the result of a 6.12% increase with respect to year 2012 [3]. Furthermore, according to the official website of the *Bureau of Transportation Statistics* of the *United States Department of Transportation*, during year 2017 pipelines transported over 18% of the freight; represented in crude petroleum, gasoline, kerosene, ethanol, diesel and other fuel oils. The value of that freight, in 2017 US

E-mail address: rjacosta@icesi.edu.co.

https://doi.org/10.1016/j.heliyon.2024.e33947
Received 22 April 2024; Received in revised form 26 June 2024; Accepted 1 July 2024

Available online 8 July 2024

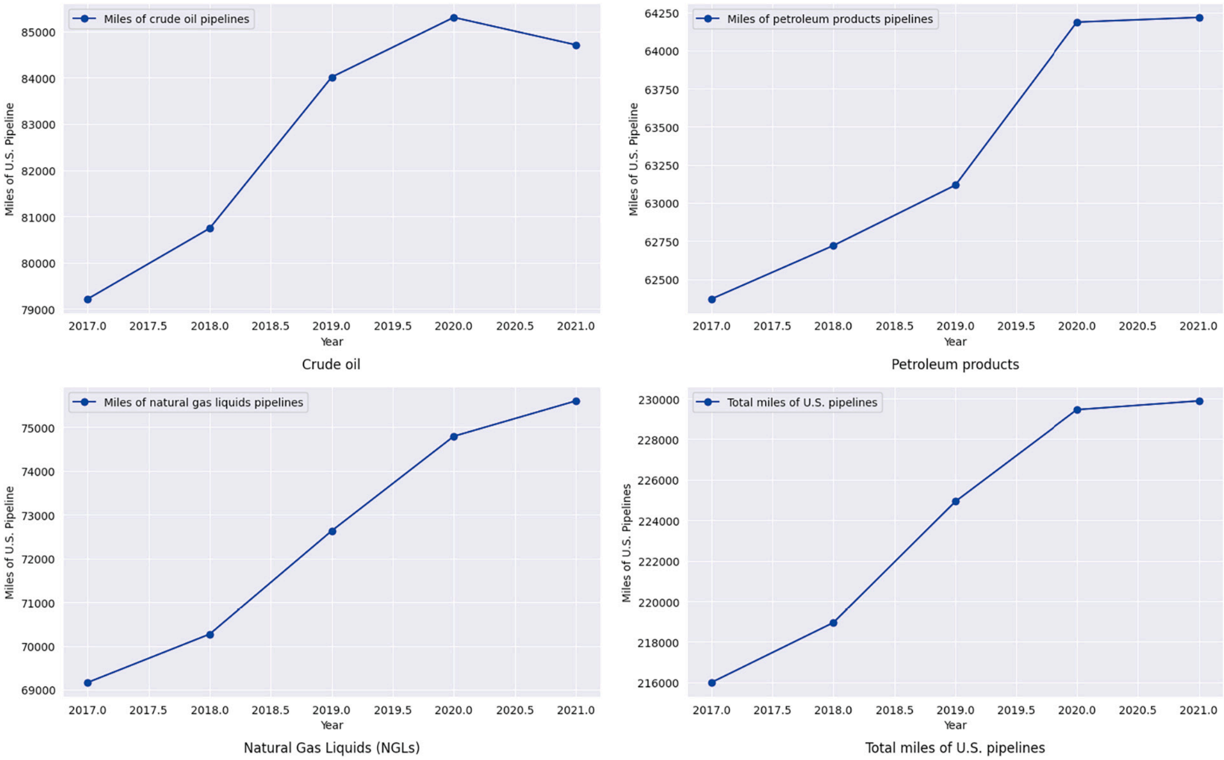


Fig. 1. U.S. pipelines mileage by transported commodities.

dollars, was over 8% of the value of all cargo moved across the USA. Moreover, two years after the beginning of the SARS-CoV-2 pandemia, the production of oil in the United States was already being forecasted to rise to record high levels during years 2022 and 2023; as then reported by the *United States Energy Information Administration* [4].

Fig. 2 illustrates a pervasive growth trend in the U.S. product supply of a) miscellaneous petroleum products, b) distillate fuel oil, c) finished motor gasoline, d) reformulated motor gasoline, e) conventional motor gasoline, f) kerosene type jet fuel, g) liquefied petroleum gases, h) aggregated crude oil and petroleum products and i) finished petroleum products. In every case, Fig. 2 reveals that, from a global perspective; product and time wise, there has been a growth trend in the number of thousand barrels supplied during the observed time lines. However, a closer look at it also reveals that there have been some points in time where an evident alteration of this growth trend has occurred. The two most recent events happened in years 2008 and 2020. In the former, an economic crisis hits the United States, and in the latter, the pandemic stressed all the economies of the world. Its effect was relatively much stronger than the effect of the economic crisis from 2008 but its recovery has also been much faster. This becomes clearly evident from the month of April in year 2020 through 2024.

The 2008 economic crisis effect is more observable in the supply of *kerosene-type jet fuel* (Fig. 2: f) as a significant decrease in the supplied volume of it is also reported; it took seven years or so to start seeing supplied volumes in quantities similar to those seen before year 2008. The 2020 SARS-CoV-2 pandemic produced a comparable effect in the supply of oil and its refined products as the one observed in the 2008 economic crisis; nonetheless, this effect progressively decreases from April 2020 through 2023.

Fig. 2 also affirms that the highest supply months have been the warmest months of the year (June, July and August) and in a similar proportion the coldest months of the year (November, December, January and February). In general, it also reveals that the trend in the supply of oil and its refined products has been *predominantly increasing* in spite of these two highly disruptive events that altered the nature of the consumption of this kind of energy.

This rapid growth of pipeline infrastructure (Fig. 1) as a response to the increasing volume of oil production and the growing volumes delivered of refined petroleum products (Fig. 2), even despite the SARS-CoV-2 pandemic, has posed an urgent need for not only safer, more reliable and efficient transportation operations of petroleum products via pipeline systems, but also robust enough to contribute to its resilience; that is, its ability to quickly reach its full operational capacity in case of a disruptive event. Also there has been a growing interest in its use for the transportation of bio fuels in Brazil [5,6] and more recently in China; where a studies centered in the use of multi product pipeline systems for the transportation of alternative forms of energy have been conducted [7,8]. Consequently, the optimization of the operation of this type of systems has been posing a challenging decision making process to the O & G sector during the last three decades or so.

In response to these facts, a significant body of literature has been published addressing different system morphology in various countries using a fairly diverse set of modeling and solution schemes. Nonetheless, a formal problem definition from an OR perspective at an operational level is still due, it remains unclear what the most influential research undertakings have been in this field, how

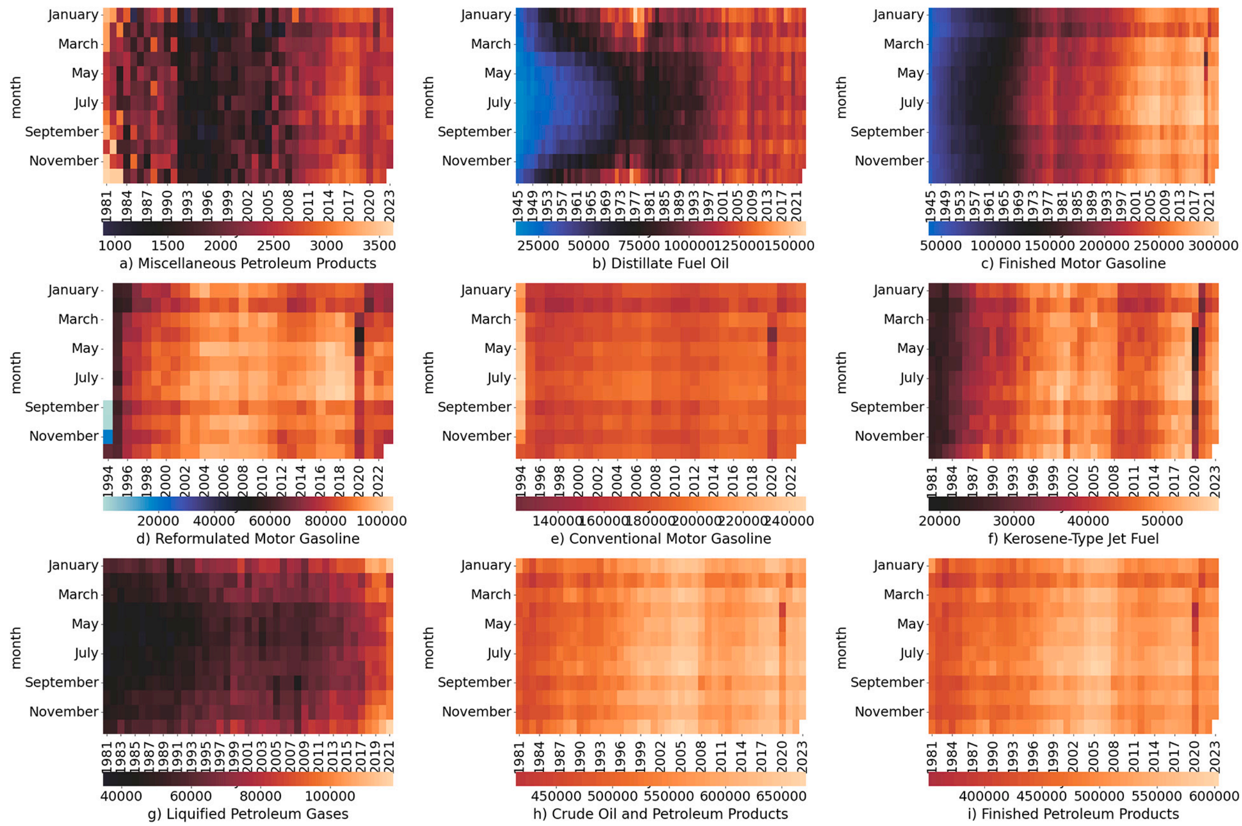


Fig. 2. Amount of oil and petroleum products supplied, in thousand barrels, in the USA across the last 30 years or more.

they relate to the rest of the body of published literature and how successful they have been to solve the problem as it occurs in realistic scenarios other than those being addressed in these studies. What is more, nor *emerging research trends* have been identified neither *a future research agenda* has been proposed; presumably, because the research progress on this topical area remains poorly surveyed.

The main contributions of this study are stated as follows:

1. **General Modeling Framework**: It presents a general modeling framework for the system from an Operations Research (OR) perspective. This framework facilitates the communication of the problem, increasing its potential for being better structured, modeled, solved, and communicated to a broader audience.
2. **Analysis and Insights**: It provides an analysis of the included research work, offering insights into its descriptive analytics and discussing research perspectives. This allows readers to better appreciate the growing importance of the Industrial Engineering (IE) and OR fields in the modeling and solution of this problem.
3. **Research Agenda**: It identifies a research agenda that outlines future research avenues for modeling and solution procedures. Addressing these avenues will require the interaction of Chemical Engineering and Operations Research, the backbone of Industrial Engineering.

The rest of the body of this paper is composed as follows: Section 2 provides a discussion on the methodology that was used to conduct this survey, which is centered in three main tasks: the population of the list of references (Section 2.1), the data preparation (Section 2.2) and the data analytics (Section 2.3). Section 3 develops two main topics: insights into the research work and its sources of publication (Section 3.1) and descriptive analytics insights (Section 3.2). Section 4 is devoted to discuss the background of the problem by addressing five frequently discussed features of it in the included literature. First, the description of the system of interest (Section 4.1). Second, the statement of the central problem (Section 4.2). Third, the type of application proposed to address the problem (Section 4.3). Fourth, the goal of the problem (Section 4.4). Fifth, the pipeline schedule (Section 4.5). Section 5 presents a discussion on the research perspectives from the most relevant literature divided in two parts: the measure of the effectiveness of a given solution (Section 5.1) and the simplifying assumptions that have been made to study this problem (Section 5.2). A discussion on the emerging and future research trends is provided in Section 6, where the modeling paradigm (Section 6.1), the challenges posed by the solution procedures (Section 6.2) and future research avenues (Section 6.3) are addressed. The conclusions of the paper are then discussed in Section 7 and then the list of references is presented at the end.

2. Methodology

This section presents a discussion on the steps of the methodology implemented in this work as well as the different tools that were used to develop the contributions announced in the previous section. The methodology is composed by two phases, phase I includes three steps: *the population of the list of references*, *the data preparation*, and *the data analytics*. Phase II provides the *analysis of results and discussion*, which contains four components: descriptive analytics of the included research work, problem background, research perspectives on this problem, and emerging and future research avenues. The aim was to include an exhaustive coverage of the research work undertaken on this problem published in *scientific articles*, *conference proceedings* and *book chapters*.

2.1. Population of the list of references

Two activities were carried out: the creation of an *early list of references* and a *saturation process*. To begin populating *a preliminary version* of the former, an electronic search was conducted on Google Scholar by using the keywords *multi product pipeline distribution optimization*, initially limiting the time span of the search between years 2015 to 2024. The combined approach of an *initial time span* and an *iterative saturation process* strengthens the findings of the study by ensuring a comprehensive review of relevant literature.

The included references are documents where the problem of scheduling a pipeline to distribute multiple oil refined products is evidently being addressed. This was determined by the title of the document, its abstract or when they included these keywords; or some other closely related in meaning descriptors. Some examples of this kind of cases included papers with any combination of the words *optimization scheduling of multi-product pipelines*, *optimal detailed scheduling of multi-product pipeline systems*, or *planning sequencing transportation product distribution pipelines*.

On the other hand, documents presenting discussions in the contexts of the supply chains, specifically those centered in the design and planning for strategic or tactical decision making perspectives, were excluded from this review. While this kind of studies might well be of a tremendous relevance for the O & G industry, there are two issues with them for the purposes of this work. First: they address the decision making process at an strategic level, and second: a broader view of the system and different decision scopes are investigated in that kind of works. One example of this type of document is the work of [9]. The only reference of this kind of undertaking that was included is the work of [10], which is highly cited across most of the research work investigated in this paper and presents an OR view of the problem.

The *saturation process* was conducted using the preliminary list of references obtained in the previous step by searching on their list of references the relevant documents and repeating this process until there was no more papers to add to the list. Some conference proceedings seemed to be shorter preliminary versions of some journal articles with either the same set of authors, or a subset of them. For this kind of cases, when the titles were the same, the conference proceeding version was excluded from the review. One example of this kind of case can be illustrated with the works of [11] and [12]; the former was excluded while the latter was included. Consequently, the list of conference proceedings was ultimately purged for the review in order to leave only those documents with significant differences from the rest of the included research work for this review. The result of this process is a list currently composed by 130 documents, out of which 108 are journal articles, 3 are book chapters and 19 are conference proceedings.

2.2. Data preparation

This part of the methodology is devoted to organizing the included references, and their information relevant for this survey. A spreadsheet was used to keep a record of the relevant data from each reference included in the review. The fields that were included in each registry of the data base were grouped in four sets depending on which the fields of each registry were related: publication, systems, model and contribution.

- publication related: *title, type of document, name of publication/venue*
- systems related: *problem definition, pipes topology, type of application (static or dynamic)*
- model related: *input data, output data, problem goal, model assumptions, time and volume domain (discrete or continuous), solution scheme, type of flow, optimization technology used, hardware used*
- contribution related: *type of objective function, elements of the objective function, modeling approach, information about the instance investigated, findings, future research, and contribution or new concepts introduced*

The classification of fields into distinct sets was integral to our methodology. Publication-related fields played a critical role in accurately identifying and cataloging each reference within our dataset. Systems-related fields were essential for synthesizing the diverse system configurations and applications discussed in the literature. Model-related fields were pivotal for analyzing the methodologies and technical aspects employed in the reviewed studies. Contribution-related fields facilitated a detailed exploration of each study's contributions to the field of multi-product pipeline distribution optimization.

2.3. Data analytics

A network of citations (NC) among the research work was developed using *The Apache PDFBox® Java library* and the Cytoscape package [13]. The Java code was used to identify the title of each paper either from its metadata or from its text, and then to find out what papers cite it in their references. The output of this code was the list of arcs of the NC; which was then exported as an excel

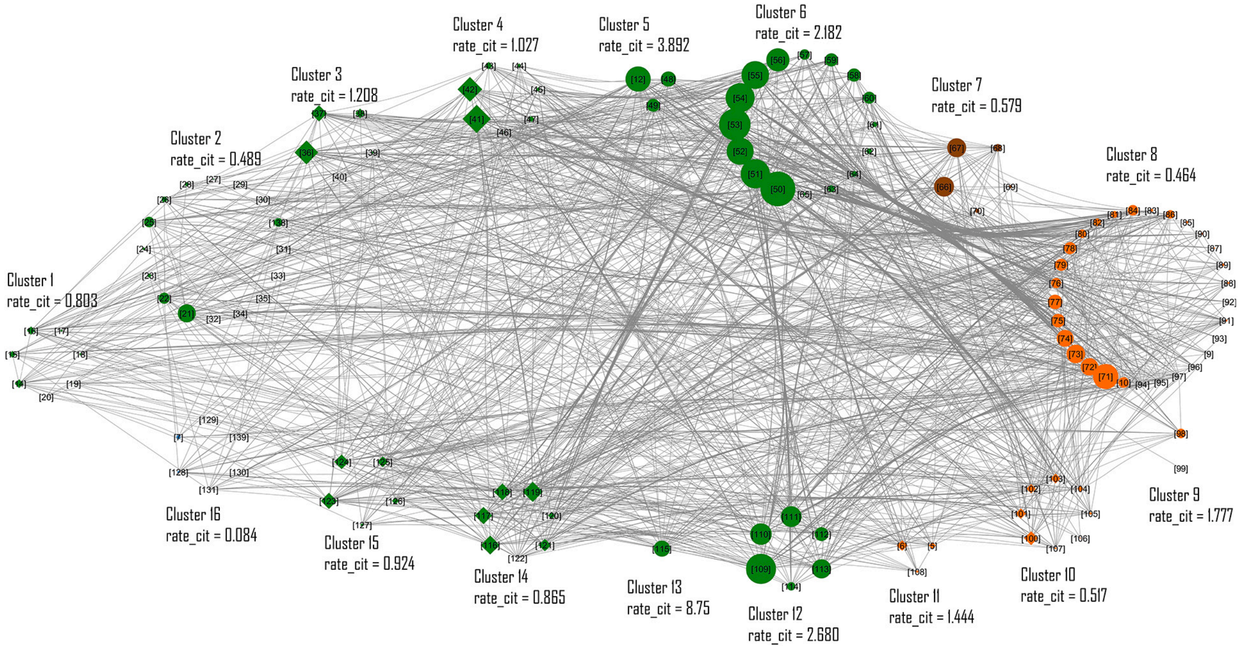


Fig. 3. Source of publication, system morphology, direction of flow.

file to create a *network from a table* in the Cytoscape package. This software, mainly used in biology to investigate huge networks with massive amounts of data, allowed the organization of the included documents in a directed network. Each vertex of the network corresponds to a reference, and the edges *(i,j)* represent a citation made in reference *j* of reference *i*. The relevance, system morphology investigated and its flow direction feature are also part of the information provided by the NC. The size of each node was associated to its relevance for the research work undertaken after its publication, the system morphology was symbolized with the shape of the node, and the flow direction capability was associated to the cluster the nodes belong to. Although several different forms to visualize the network were investigated, only one of them is reported in this paper due to space limitation. The NC was analyzed after clustering its nodes by source of publication and then considering two system features: its morphology and the ability to perform reverse flow operations. Moreover, the illustrations for the analysis of the included research work were created using either Matlab, Matplolib or Seaborn in Python, with the aim of unveiling the shape of the investigated research work from different perspectives.

3. Descriptive analytics on the included research work

3.1. Insights into the research work and its sources of publication

The process described in section 2.3 not only arrived at insights from the descriptive analytics on the preferred journals and venues for publications among the authors in this topic, and how they have interacted; but also it facilitated the identification of the most influential research work on this topical area. It also helped revealing emerging research trends and highlighted interesting relationships among the different documents. These all are not obvious insights when just reading the papers separately without considering their interactions. The discussion of this section is about the current picture of the research work included in this survey without considering its evolution through time to arrive at this state.

Fig. 3 displays 16 circle clusters for the research work organized by the number of citations clockwise in decreasing order and Table 1 provides the list of references included in each cluster. The reference with the smallest number of citations in every cluster from Fig. 3 is at the bottom of the circle, while the paper with the largest number of citations corresponds to its left hand side neighbor; which itself is the first of each list presented in Table 1. The sizes of the nodes represent the *mean citations per year*, which is obtained by dividing the number of times the document has been cited by the number of years ever since its publication. There was not a case in which a citation of a work was made to debunk a wrongly conceived part of it, such as a set of assumptions, a model, solution procedure or a set of results or conclusions. Consequently, it is reasonable to consider every citation made of any of the included references as a measure of their relevance in a positive sense.

The domains of the sources of the publications were classified in: *Chemical Engineering (CE)* (green nodes), *Industrial Engineering, Operations Research and Applied Mathematics (IEOR-AM)* (orange), *Petroleum Science (PetSci)* (brown nodes) and other sources of publication from *Miscellaneous Engineering (MiscEng)* (blue nodes).

Table 2 explains the abbreviations from Table 3, which provides additional information of the research work at each cluster. Such information includes the domain of the publications, their source, the morphology of the systems investigated, and the type of flow operations the system can perform.

Table 1
Cluster composition in Fig. 3.

Cluster	References
1	[14–20]
2	[21–26,28,27,29–36]
3	[37–41]
4	[42–48]
5	[12,49,50]
6	[51–66]
7	[67–71]
8	[10,72–94,9,95–98]
9	[99,100]
10	[101–108]
11	[6,5,109]
12	[110–115]
13	[116]
14	[117–123]
15	[124–128]
16	[129,7,130–133]

Table 2
Abbreviations from Table 3 and their descriptions.

Abbr.	Description
CE	Chemical Engineering
PetSci	Petroleum Science
IEOR-AM	Industrial Engineering, Operations Research and Applied Mathematics
MiscEng	Miscellaneous Engineering
N	Network of pipelines
S	Single pipelines
U	Unidirectional flow operations
B	Bidirectional flow operations

Table 3
Domain, morphology and flow features in Fig. 3.

Cluster	Domain	Source(s)	Morphology	Flow
1	CE	Miscelaneous	N	U/B
2	CE	Miscelaneous	S	U/B
3/4/5/6	CE	*Computers & Chemical Engineering*	N/N/S/S	B/U/B/U
7	PetSci	Diverse related journals	S	U
8/9/10/11	IEOR-AM	Diverse related journals	S/S/N/N	U/B/U/B
12/13/14/15	CE	*Industrial & Engineering Chemistry Research*	S/S/N/N	U/B/U/B
16	MiscEng	Miscellaneous Journals	S,N	U,B

Let *rate_cit* be a ratio defined for each cluster to represent the total number of citations that have been made of all the documents within that cluster divided by the total number of works cited by those documents. This ratio provides insight into the scope of the contribution of these documents to the literature on this problem. For every cluster, the corresponding *rate_cit* appears under its name in Fig. 3.

According to this ratio, in the CE realm, studies centered on single pipeline systems with bidirectional flow capabilities (clusters 5 and 13) have made quite a significant contribution to the literature; perhaps due to the realism and level of modeling complexity involved in the bidirectional flow operations. Although they have been outnumbered by far by their unidirectional counterparts (clusters 6 and 12), fairly interestingly, at the same time, those documents have been taken into account more often by the subsequent studies on this problem than their unidirectional counterparts. Notwithstanding, the single unidirectional pipeline studies have also made a significant contribution to this field. Moreover, although the number of papers on this version of the problem in cluster 12 is significantly lower than in cluster 6, the former has been relatively more influential in the subsequent literature than the latter. Recall that papers in cluster 12 were published in the *Industrial & Engineering Chemistry Research* journal and those in cluster 6 in the *Computers & Chemical Engineering* journal.

The works from clusters 3, 11 and 15, are all centered in systems with network morphology with bidirectional flow, with a relatively small number of works. The level of complexity of the system investigated in those documents is remarkably higher than the case of a single unidirectional pipeline scenario. The works of cluster 11 were published in the *Principles and Practice of Constraints Programming* journal and the *Constraints* journal. Centered in a more manageable scenario, the clusters that follow according to their *rate_cit* are 4 and 14. These all investigate systems with network morphology but no bidirectional flow capability on their pipelines; which still poses a significant challenge from the modeling and solution procedure perspective. The works from clusters 8 and 10,

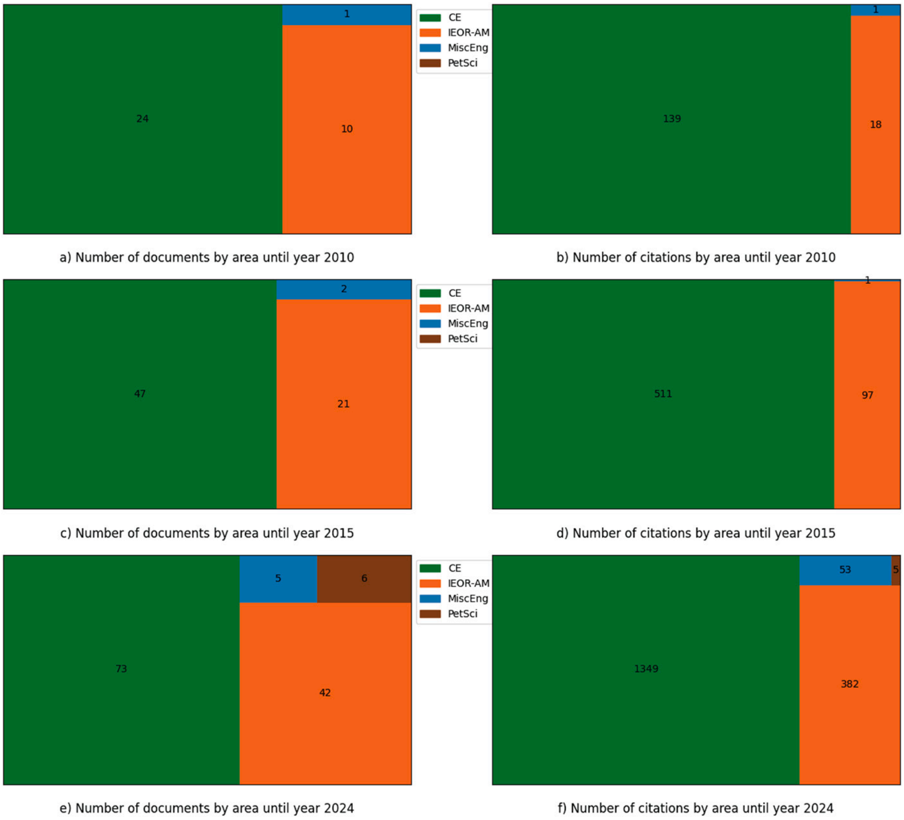


Fig. 4. Publications, citations and relevance.

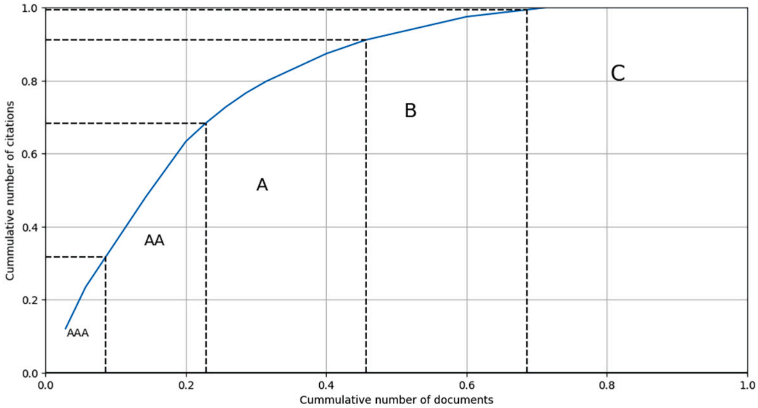
which are all centered in single unidirectional pipeline systems, and were published in sources from the IEOR-AM sphere, seem to have a relatively much smaller *rate_cite* when compared with all the previously mentioned clusters.

A view on the arcs of the network suggests that the strongest connections have been forming with the course of time among clusters 6, 8 and 12; all of them composed by works addressing systems with a unidirectional single pipeline; some of them published in the *Computers & Chemical Engineering* journal, journals from the IEOR-AM sphere and the *Industrial & Engineering Chemistry Research* journal. Among these three, the latter is the smallest in both number of references and citations, while in contrast there seems to be a stronger connection among the works of the first two. While 444 citations were identified for the works published in *Industrial & Engineering Chemistry Research*, 771 citations were counted for the included research work published in *Computers & Chemical Engineering*. The citations of the works published in this journal addressing systems with single unidirectional pipelines (*Cluster 6*) are 467; greater by itself than the number of citations of every study from the *Industrial & Engineering Chemistry Research*. In less proportion, still significant, clusters 3 to 5, 9 to 11 and 13 to 15 complete the vast majority of the rest of the connections. On the other hand, 374 citations were identified for the research work in clusters 8 to 11, documents published by sources in the IEOR-AM domain, of which 243 are from undertakings addressing systems with unidirectional single pipelines morphology, a predominant portion of the total.

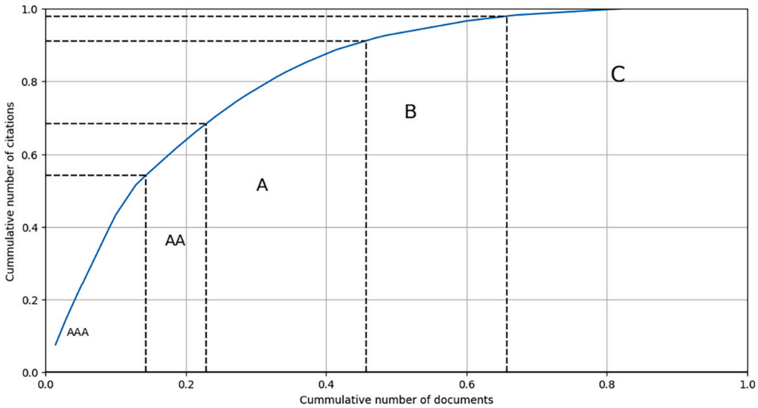
3.2. Descriptive analytics insights

The total number of references included in the analysis is 130. However, not all the references were considered in some discussions when the necessary data was not available. Fig. 4 illustrates the progression of research work at three distinct points in time, highlighting its historical evolution through the number of publications and citations. The treemaps are for *the total number of documents published in each domain* and *the total number of citations identified of the documents in each domain*. Fig. 5 presents a Pareto plot of the number of citations. The three time lines for the analysis of the included research work are years 2010 (Fig. 4 a, b, Fig. 5 a), 2015 (Fig. 4 c, d, Fig. 5 b), and 2024 (Fig. 4 e, f, Fig. 5 c).

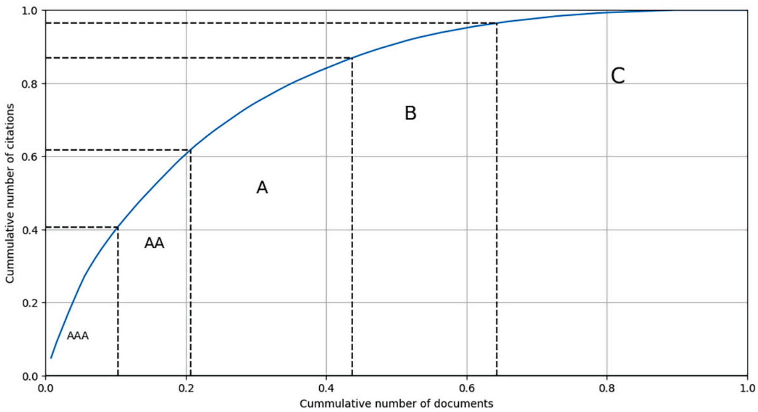
Fig. 4 *e)* illustrates the number of publications from each of these four areas up until year 2024 revealing that more than half of them are in the CE sphere, while about a third of the research papers are in the IEOR-AM realm. The number of publications from the PetSci orbit was equal to the number of publications from other miscellaneous engineering (journals or venues). This suggests that the research activity on this topical area has occurred in more than 90% of the cases in the CE and IEOR-AM domains. Furthermore, for every 20 documents that have been published on this problem, there are roughly 11 from the CE field, seven from the IEOR-AM, one from PetSci and one from other engineering related sources of publication. Quite interestingly, these proportions change when the



a) Pareto plot of citations per documents until year 2010



b) Pareto plot of citations per documents until year 2015



c) Pareto plot of citations per documents until year 2024

Fig. 5. Citations vs. Publications Paretto chart.

number of citations up until this point in time are considered. About three out of every four citations have been made of a document from the CE field, one out of every five citations have been made of a document from the IEOR-AM field, two out of every 10 citations have been made of a document from the MiscEng domain and only two out of 100 citations have been made of a document from the PetSci domain. The number of publications, as well as citations, of the research work from the PetSci area is somehow unexpected as the system where this problem occurs belongs to the downstream supply chain in the petroleum industry. In contrast, the CE realm has been the major contributor to this problem in both number of publications and relevance.

A total of 1789 citations among 126 of the 130 documents included in this survey were identified. A remark worth noting is that the *proportion of citations* of the research work from the CE field, as seen up until year 2024 (Fig. 4: f), was *noticeably bigger* than the *proportion of publications* from the same field (Fig. 4: e). Contrary to this, up until the same point in time, the proportion of the citations of the included research work from the IEOR-AM area was significantly *lesser* than its publications counterpart. Nonetheless, seeing this without considering the past can be misleading as far as the conclusions on the role of the IEOR-AM realm are concerned. This discussion must be completed with a review of the state of the research work up until years 2010 (Fig. 4: a, b) and 2015 (Fig. 4: c, d). The whole picture reveals that both the number of publications, and their relevance, *from the IEOR-AM sphere have been increasing*. Both, the number of published documents on this topical area from the IEOR-AM domain as well as the number of citations of these undertakings, have been gaining participation on the scientific literature of this problem.

Fig. 5 (a, b, c) provide a Pareto classification of the research work included in this survey; the same criteria was used to determine the cut points for each class in every one of them. Five sets of works were identified based on their citation counts, ranked from the most cited to the least cited: AAA, AA, A, B, and C. The division of the most cited works further distinguishes those with a high level of citations (AAA) and those with moderate levels (AA and A). A ratio between the *cumulative number of citations* to the *cumulative number of publications* was calculated and the cut points to separate the classes were determined as 4.0, 3.0, 2.0 and 1.5. References with a ratio greater than or equal to four are in class AAA, references with a ratio less than four and greater than or equal to three are in class AA and so forth.

Fig. 5 (a, b, c) illustrates an increasing separation between papers with higher and lower citation counts, with fewer papers in the middle categories. Over time, the sets of the most cited works AAA, AA, A have fluctuated from 42% in 2010, 44% in 2015 and back to 42% in 2024; roughly the same proportion over time. In this group of works, the set A has constantly represented the 22% of the literature through time. On the other hand, there is a fluctuation in the size of the set AAA. Up until the year 2015, the size increased, representing roughly 12.8% of the references, compared to the year 2010, when it was 5.7% of the references. Some references published before 2010 solidified its influence in the field during this period [51,12,52,54,53,55,110,10,116]. By year 2024, this percentage was 9.6%, which suggest that the field is becoming more selective as fewer papers have been receiving a moderate to high number of citations. As of 2024, the size of the AAA set had decreased, but it remained larger than its size in 2010. This trend reveals two plausible facts: first, although the scientific work on this topical area has continued to develop, the relevance of the published work has been increasing at a lower speed. Second, despite the set of the most relevant publications has tapered off with the course of time, some papers have been consolidating their lasting recognition. Despite the increasing complexity of the problem, a breakthrough in modeling and solution procedures in this field remains outstanding.

On the contrary, the size of the C class has been consistently growing, which signifies that there is a mounting number of papers with low citation counts, or not widely recognized in the field. It also suggests that many papers are not gaining traction or recognition in the scholarly community. As time progresses, there is a clearer separation between highly cited and poorly cited papers. This indicates a polarized citation landscape where papers are either highly influential or largely overlooked. This translates to a few papers driving the research agenda by receiving most of the citations and influencing future research directions. On the other hand, many papers might not be receiving the recognition they deserve, even though they have addressed systems under challenging conditions for both modeling and solution procedures.

4. Problem background

This section of the paper is divided in five parts addressing the system of interest, problem statement, type of applications, problem goal, and the pipeline schedule. The first part is devoted to discuss the system in which this problem occurs as studied in the literature. The second part aims to synthesize the problem statements presented in the literature into a general definition from an OR perspective. The third part proposes a classification of the type of applications and provides comments on their conveniences and computational implications. The fourth part identifies the most common five goals when addressing this problem as found in the included literature. The fifth part provides insights into the pipeline schedule according to the main authors from this topical area.

4.1. System of interest

This section of the paper introduces an abstraction scheme that can be used to represent any of the system morphology addressed by the included references in this survey; which can vary from a relatively simple *single source and single destination connected by a one segment pipeline system* to a much more complex scenario where *multiple sources and destinations* are connected among themselves by a set of pipelines somehow disposed in a sort of *network shaped* system of distribution.

Without loss of generality, the system of interest where this problem occurs is composed of three kinds of elements: *locations, pipelines* and *refined products*. A *capacitated* network scheme to model these elements is proposed. The locations are represented by the set of nodes $\mathscr{L}$, the set of pipes of the distribution system are represented by the set of directed arcs $\mathscr{A}$ and the capacity of each arc represents the volume capacity of the pipe it corresponds to. The products transported by the system are represented by the set of commodities $\mathscr{C}$ that flow through the network. This arrives at the representation of the system by the directed network [134] $\mathscr{N} = (\mathscr{L}, \mathscr{A})$ and the set of commodities $\mathscr{C}$ to be distributed.

From the perspective of their role with the products being transported, three types of locations with different features were found in the included literature for this survey:

1. The source nodes (refineries, ports, etc.), which represent the locations of the system where the flow of products is originated. These nodes are grouped in the set Ω.
2. The intermediate nodes, which have the ability to work as both source and destination nodes. These nodes represent dual locations and are grouped in the set $\mathscr{I}$.

3. The destination only nodes (market zones, final clients), which represent the locations of the system where the flow of products is consumed through the satisfaction of their demand at these points. The destination only nodes are grouped in the set Δ.

The set $\mathscr{L}$ results from the union of all these three sets: $\Omega, \mathscr{I}$, and Δ as $\mathscr{L} = \{\Omega, \mathscr{I}, \Delta\}$; which are disjoint sets.

The pipelines through which the different batches of products traverse the system can be either unidirectional or bidirectional, and as previously mentioned, they are represented with the elements from set $\mathscr{A}$ as the arcs linking the elements from the set $\mathscr{L}$. However, these arcs do not operate likewise the arcs of a plain network flow model do. Their capacity represents the volume of liquid that can occupy the pipe it models and so all the arcs must be used at full capacity all the time. The included research has used a set of assumptions made on the pipelines of the system, discussed in Section 5.2.3, to specify the unique dynamic properties of the system. These assumptions are an approximation to the dynamic details of the operation of the real system. Finally, the oil refined products that are transported through the system, which have included bio fuels in previous undertakings as well, are represented by the elements from set $\mathscr{C}$, which are the commodities that flow through the network $\mathscr{N}$.

The arrangement of these three sets of different elements (locations, pipes and refined products) poses a wide diversity of possibilities for the morphology of the system, which results in different levels of system complexity, and therefore, different levels of challenge in modeling and solution procedure endeavors.

4.1.1. Nature on the divisibility of goods

According to [51], poly-ducts operate in two different modes: *batch* or *fungible*. Batch driven operations preserve the identity of the product shipped through the system until its delivery occurs: the same material that was accepted for transportation at the origin is actually delivered at its assigned destination. In contrast, fungible mode operations are more flexible in this dimension as the system does not exactly deliver the same batch of product that was dispatched at the injection point. Rather, the system may deliver a volume of product different than the one originally injected into it, of course keeping the same specifications at its destination. This kind of operation is possible when the system carries generic products and has operational conveniences as it minimizes the generation of product contamination, allows batch stripping operations downstream the lines and promotes a more efficient use of storage capacity; which also might positively impact inventory management policies. As a result, in a fungible operational mode, a batch of product can be allocated to two or more terminals by entering a portion of it into one of the assigned destinations while the rest of it continues downstream the line towards the rest of the delivery points. This kind of operation is known as *batch stripping* [55]. In this context, goods can be either non divisible if the system operates in batch mode or divisible if the system operates in fungible mode.

4.2. Problem statement

The aim of pipeline logistics is to ensure that the right product will be made available in the premises of the customer on the right time and at the lowest cost under the best possible conditions of safety, security and respect for the environment [43]. One of the pioneering works in this topical area defined the problem as "*the optimization of the transportation of various grades of refined products through a system of pipes*" [10]. The model proposed in this work sought the minimization of a *steady state sequence* dependent costs of delivering all the orders in a *static order* sequence. Optimization work was supposed to be done neither on the pumping sequence nor on the delivery operations. Instead, the measure of effectiveness was related with the reduction of the variance in energy demand, as a result of the minimization of the number of stoppages. As mentioned in Section 2.1, in contrast with the rest of the literature on this problem that addresses it from an operational decision making perspective, this paper aimed to respond to strategic decision making level concerns.

On the other hand, from an operational point of view and without loss of generality, the problem has been defined in the literature included in this survey as the optimization of the operation of a system of distribution of refined products from a set of supply locations towards a set of consumption locations to satisfy their demand for these refined products. The products are all in a liquid state and they must travel long distances before reaching their destination. Pipelines are therefore used and the transportation operation must be performed with the best possible measure of effectiveness.

The solution of the problem provides "*the optimal sequences of products pumping, the volume of each batch of product, the destination of each batch of product and the sequence of discharging operations at the destinations downstream the line*" in order to satisfy every product demand at each depot while meeting time requirements [92]. The optimal product sequence and the corresponding sizes of those lots, involve the solution of a combinatorial optimization problem that should minimize the product contamination costs while the promised delivery dates are guaranteed [87]. Furthermore, it is also required to keep the inventory level of every product at every location in the system between the corresponding permissible ranges all the time.

In line with this, for [57] the objective of the problem is to find the optimal volumes and timing of product injections from input nodes and deliveries to output nodes that meet demand (assumed to occur at the end of the time horizon) in the shortest time possible, while respecting flow rate bounds and forbidden product sequences. What is more, when a volume of contaminated product is created, as a result of sequencing two non miscible products, it has to be sent back to the refinery to be separated, which results in a significant increase in the operating costs of the system [92]. Therefore, a pipeline schedule should not place two non miscible products consecutively.

In the case of systems with network morphology, the problem definition has included the determination of the pipeline route each batch must follow to complete its journey through the distribution system [109,5,127]. For [6], a pipeline in a route is called a segment and a route is an alternating sequence of depots and *non repeating* connecting pipelines. However, if the system performs reverse flow operations, then the latter is not true. As a result, all products being transported by the system must have a well defined route to follow in order to move from its origin node towards its destination node [6]. According to [41], there is a finite set of routes when the system is represented as a graph $\Gamma(N, A)$ where N is the set of nodes and A is the set of arcs. The former represents the locations of the system and the latter represents its pipelines. Based on that modeling scheme, a route is then defined as a sequence of arcs from A and forms a valid path between two types of locations: a source, and a destination. Under this scheme, however, nor the intermediate dual purpose locations are well identified neither their characteristics well defined.

In simple topology systems the path of pipelines traveled by the batches of product can be just one pipeline segment between the origin and the destination, yet when more complex scenarios are investigated, the options to define such path grow significantly. Furthermore, the pipelines might have segments with different diameters and in some cases, they might also perform bidirectional flow operations. Unless specified otherwise, the destination nodes must satisfy the requirements determined by local consumer markets and there is no physical separation between two consecutive batches of different products [33,29,36,52,53,51,12,27,135,110,55,54,43,85,35,92,42,89,56].

Let Π be the set of batches of the products in $\mathscr{C}$ available among the elements of the set Ω. Let $k \in \mathscr{C}$ be a product to be distributed by the system of which there is a set $B_k \subseteq \Pi$ of batches available among the elements of the set $\omega_k \in \Omega$ that will be used to satisfy its demand at the elements of set $\delta_k \subseteq \Delta$. Let $\beta_{k,\lambda,o} \in B_k$ be the batch λ of product k leaving from its origin $o \in \omega_k$. A set of destinations $\alpha_{k,\lambda,o} \subseteq \delta_k$ must be determined where their demand for product k shall be fully or partially satisfied using the product content in the batch $\beta_{k,\lambda,o} \in B_k$.

Definition 1. A path of pipelines $\rho_{k,\lambda} \subseteq \mathscr{A}$ is an ordered set of pipeline segments represented by the set of arcs $(i,j) \in \mathscr{A}$ through which the batch $\beta_{k,\lambda,o} \in B_k$ must travel downstream the line from its origin $o \in \omega_k$; once it is injected into the pipeline system, to visit each member of its set of destinations $\alpha_{k,\lambda,o} \subseteq \delta_k$ in order to fully or partially satisfy their demand until it is totally siphoned from the system.

The *path of pipelines* for every batch of product is the backbone of its *schedule of distribution*. Furthermore, the definition of the latter for every batch of any product transported by the system is an important part of the solution procedure for the problem as well.

Definition 2. The schedule of distribution ζ_{β_k} for any given batch $\beta_{k,\lambda,o} \in B_k$ is defined by its volume V_k, a path $\rho_{k,\lambda} \subseteq \mathscr{A}$ of pipeline segments it must traverse, a set $\alpha_{k,\lambda,o} \subseteq \delta_k$ of destinations downstream the line it will serve, and for each of these destinations, a volume $v_{d,k}, d \in \alpha_k$ of it that must be discharged from the line towards the satisfaction of their requirements of the product from the batch.

The problem solution must then provide, for every product $k \in \mathscr{C}$, the *schedule of distribution* for every batch $\beta_{k,\lambda,o} \in B_k$ in the sequence to be pumped from its origin $o \in \omega_k$ towards a determined set of destinations $\alpha_{k,\lambda,o} \subseteq \delta_k$, $\delta_k \in \Delta$.

Definition 3. The pattern of distribution for the set of batches Π of products in $\mathscr{C}$ available at the elements of the set Ω and demanded at the elements of the set Δ is composed by the schedule of distribution for every batch $\beta_{k,\lambda,o} \in B_k$, $B_k \in \Pi$.

Formally, and given the previously defined elements, the problem of optimizing the transportation of refined products using pipeline systems can be stated as follows:

Definition 4. Let $\mathscr{N} = (\mathscr{L}, \mathscr{A})$ be a directed network where liquid commodities from the set $\mathscr{C}$ must be transported from the origins in the set $\Omega \in \mathscr{L}$ to satisfy their demand at the destinations in the set $\Delta \in \mathscr{L}$, passing through a set of intermediate points in $\mathscr{I} \in L$, using the arcs from the set $\mathscr{A}$, which must be full of commodities all the time, pushing each other to move forward or backward without any empty space separating them, find the pattern of distribution with the best possible measure of effectiveness.

The *pattern of distribution*, from Definition 3, for the refined products to follow in order to fulfill their demand at the consumption nodes while optimizing a given measure of effectiveness is the solution for the problem.

4.3. Type of applications

According to the availability of the demand data to address the problem, two types of applications were identified in the included literature for this survey: *static* and *dynamic*. In static applications, demand data are available only once for the horizon planning period and the schedule has to be generated for that period of time based on that data. Consequently, once the problem is solved no later modifications can be made to the solution identified. Pipeline schedules generated for short term periods often do not reflect customer requirements ahead due to the operational lag [43]. In contrast, in dynamic applications new demand data are included as the planning horizon is in progress to update the optimal solution from the previous version allowing the consideration of longer time horizons [27,116,55,35,92,129,91,80]. In this kind of applications, the sequence and volumes of new product batches to be pumped in the pipline are *dynamically updated* throughout the multiperiod rolling horizon. This modeling and solution scheme has the ability to deal with longer time horizons and has resulted in a more realistic view of the problems and pumping schedules of better quality [92,91]. Nonetheless, a limitation worth noting is the high computational burden that apparently increases much faster than the size of the system under consideration. This setback on this kind of applications could be addressed with the development of more computationally efficient solution procedures.

4.4. Problem goal

There has been a significant variety of goals when addressing this problem with a fairly diverse set of levels of complexity from the optimization point of view. For example, some authors state their goal as *to determine a feasible solution for the problem* [109,6,44,58,76,45,68,67,60] while other references pursue the optimization of an operational measure of effectiveness related with time [30,20], customer demands [127,114,57,77,104,28]; or an economic criterion resulting from the optimization of pumping schedules, which is an operational decision [70,16,79,64] or a combination of this with customer demands [112,114,74,133,72,61,21]. Nonetheless, some other works seek the determination of not just a feasible schedule, but optimal as well, considering a set of economic, operational, or both, criteria [111,124].

Without loss of generality, in either *static or dynamic* applications, the goal when addressing the problem has been fivefold: to minimize a cost function, which has a fairly diverse set of criteria [130,27,111,38,43,117,92,112,37,73,19,89,50,75,103,78,132,74,72,82,133,129,106,91,21,80,61,16,105,64,65,93,86], to meet every product demand at every terminal in a timely fashion [27,5,111,38,92,37,127,73,114,50,103,74,83,72,82,133,125,57,106,77,61,105,26,93,86], to maintain the inventory level in the refinery and terminal tanks within the permissible ranges [27,38,92,41,76,83,106,60,105,26,65], to enforce the flow rate to be within the permissible ranges along the line [73,114,56,125,88,105,65], and to provide the required data to trace the batches in transit to facilitate the monitoring of their size and location inside the line, which has been explicitly mentioned in a few works [117,37,73,19]. The latter is of central importance to provide the system with the ability to respond to conditions as those generated by disruptive unplanned stoppages.

4.5. The pipeline schedule (PS)

According to [51,110,55,49,32,34], the PS defines the sequence of products to transport, and for each product, the volumes of the batches of product, pumping rates and the associated timing requirements, such as the beginning and ending times of each lot pumping and discharging operations over the entire horizon planning period. The pipeline stoppages during the horizon planning period must also be identified [30,32]. The PS enables the inventory management at the distribution centers during the period of interest. This includes the daily volume balances by product and the monitoring of arrivals, the setting and approving tasks as well as the satisfaction of the demand of the clients for these products.

The PS must indicate the *amount and type of product to be pumped, the batch pumping rate as well as the starting and completion time of every batch injection* [27,55]. Besides these, the PS shall provide the details of the decision making process related with the inventory management at the distribution centers; including daily volume balances by product and monitoring of arrivals, as well as the settling and approving tasks and the fulfillment of demands. All these details must be oriented towards the optimization of a given measure of effectiveness [24,110,55,49,32,34,76].

Under those schemes, a push type of policy to operate the system is adopted: first, how the products are to be sent is decided; and somehow, as a side effect, the inventory policies must be identified, in second place. Both decision making processes should actually be part of the same model in some sort of hierarchical managerial process that uses a global measure of effectiveness. This objective function should be connected to both the inventory management and the pipeline scheduling processes.

Furthermore, flow rates of each segment, locations of different batch interfaces and densities of the refined product flowing through pump stations may also be included as part of the PS [64]. What is more, in the case of systems with a network morphology, the PS must specify for each pumping operation not only the product, volume and timing for each batch injected into the system, but also their *routes across the network of pipelines*. This piece of information shall include of course the origin and the destination tanks [6].

The pipeline scheduling process is aimed at developing both the input and the delivery schedules. The input schedule indicates the pumping runs at every input terminal, as well as the injected product, batch size, starting time and pump rate for each run. The delivery schedule specifies the product batches partially/totally leaving the pipeline and the amounts diverted to the assigned destinations on every pumping run. In addition, it provides the times at which pumps should be turned on/off and valves at terminals are to be open/closed for accomplishing the delivery plan. Its main goal is to meet depot demands while lowering the number of pipeline stoppages and pump switching to get savings on both energy costs for restarting flow in idle segments, and pump maintenance costs [87].

5. Research perspectives from the most relevant literature

The included literature in this survey provides discussions on a wide variety of problem features with different scopes. Nonetheless, for the sake of space, this section of the paper is centered in two main topics as found in the included research work: the objective function (Section 5.1) and the model assumptions (Section 5.2). The measure of effectiveness of the PS that results from the decision making process involved in this problem provides a view of what has been sought in previous studies. Moreover, the conditions under which the real system should be operating according to the literature portrays the degree of realism that has been involved in those undertakings. In both cases, the discussion addresses the most common elements an interested reader could come across aiming to highlight their importance. In some cases, a characterization of those features is proposed.

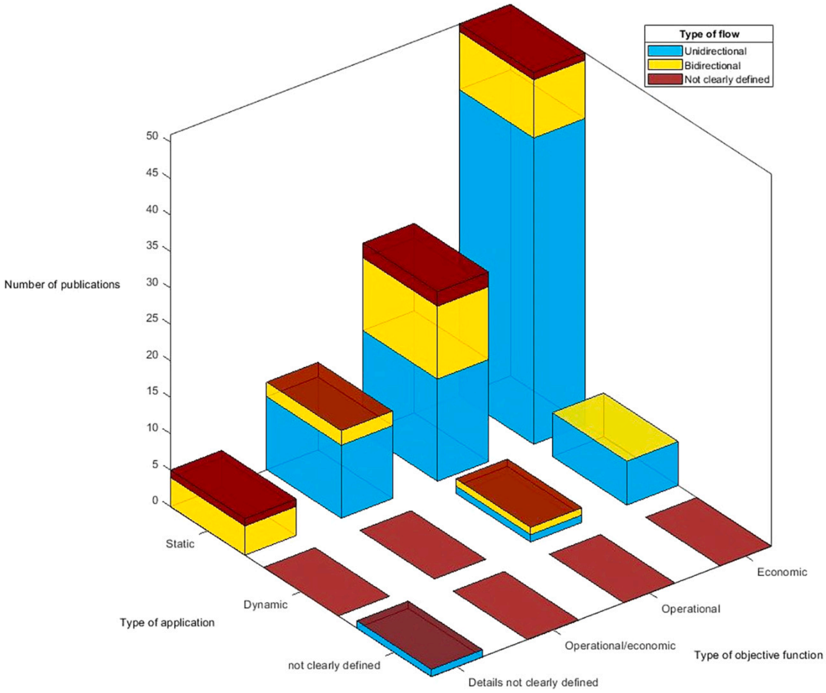


Fig. 6. Type of application, type of objective function and type of flow.

5.1. *Objective function*

There have been several measures of effectiveness associated to a pumping schedule: some of them economic, some others operational and in less proportion, a combination of these two. Fig. 6 reveals that most of the research work undertaken on this problem has been primarily centered in static applications with economically driven measures of effectiveness for the most part, and in much less proportion, operational objectives. In contrast, the dynamic version of the problem with economic objective functions has received a little attention and even less interest has attracted the dynamic applications for the problem with operational measures of effectiveness.

12 different measures for the effectiveness for a pumping schedule were found in the included literature, which can be further classified into two categories: economic and operational. Three of these objective function criteria have been used in roughly 80% of the research work included in this survey and all of them are economic. Ordered by frequency of use, from largest to smallest, these are costs related with the *energy consumption*, the management of *product inventories* in the different locations of the system, and the *processing of the contaminated products*. The rest of the measures of effectiveness is composed of measures of operational performance, such as the make-span of the scheduling horizon, the amount of product pumped, the capacity under utilization, the number of cut operations, the deviation from the mean rate of flow, the idle volume, the product changes and the number of pipes requiring reverse flow operations. [56] also includes the fixed costs of operation as part of the optimization criteria.

5.1.1. *Energy consumption*

One of the seminal works on this topical area stated that "*for an application to be considered realistic, the effect of the friction among the products and the pipelines in the energy required to pump them downstream the line should be well represented in the form of pumping costs*" [10]. According to [135], the pumping costs are composed by the overall power consumption of the booster stations and the time interval that the pipeline remains in operation. The unit pumping cost is a known constant that varies with the product and the stated destination, yet independent of the pumping rate [55]. Pumping costs are proportional to the amount of each product sent by the refinery and to the distance it covers in the pipeline until it reaches its destination depot [94]. These costs are associated to the energy consumption related with the pumping operations [82,67,21,61,16]. The pumping cost is determined by multiplying the unit cost of transporting a product from an origin to a destination by the volume of product transported and then added across each time slot [46]. For [38], this cost depends on the connection and the product and its minimization is achieved by minimizing the amount of product pumped. In the work of [136], the pumping cost is given by the electricity energy used by the pumps. In two later studies, [92,39] two types of these costs were considered according to the time of the day the pumping operations are performed (normal and peak hours). What is more, the value of pumping costs may vary with the pipeline segment and the product [39]. Table 4 presents a detailed classification of the various types of costs associated with pipeline operations, summarizing how each type has been considered in the literature.

Consideration of variable flows in the system, as well as the representation of related costs, add another layer of both realism and complexity; such approach would require nonlinear expressions. This model feature poses a significant challenge that demands

Table 4
Classification and summary of energy consumption costs.

Cost Type	References	Notes
Energy consumption costs	[118,70,76,50,83,67,125,21,16,64]	Also called electricity costs
Switching costs	[136,119,112,87,19,114,74,120,75,72,82,105,86]	Stop and start (on/off) costs
Pressure related costs	[130,32,89,56]	Cost of maintaining the pipeline pressurized and idle
Flow reversal costs	[50]	High additional costs due to pump and valve bidirectional flow operations
Unchanging energy costs	[117,42,73,75,78,71]	Assumed constant for each product; unaffected by the pumping rate
Injection point dependent costs	[73]	Varies with the injection point
Flow rate dependent costs	[89]	Dependent on flow rate
Pumping costs by product and connection	[38]	Cost depends on the connection and the product; minimization achieved by minimizing the amount of product pumped
Electricity energy costs for pumps	[136]	Pumping cost given by the electricity energy used by the pumps
Pumping costs by time of day	[92,39]	Two types of costs according to the time of the day (normal and peak hours)
Segment and product dependent costs	[39]	Costs may vary with the pipeline segment and the product

more research efforts to be done in both modeling and solution procedure avenues. Additionally, pumping costs have been considered independent of the product pumped [44]. Furthermore other works have assumed pumping costs to be directly proportional to the pumped volume, dependent on the product injected and on the direction of flow [50].

More recently, the pumping costs have been further classified in two categories: costs of running the pumps and costs of starting and stopping the pumps [64]. Moreover, another quite recent study [86] considered the cost of operating the pumps to be composed by the costs of flow restarts, flow stoppages, and on/off pumping switching.

5.1.2. Product contamination costs

[92] highlights additional features of astounding importance in the problem definition that must be mentioned in this section as they are considered either explicitly or tacitly in the vast majority of the research work on this topical area. First, the absence of a physical separation among the different batches of products as they travel through the line results in what is called *product contamination*, also known as *trans-mixes* or *interfaces*. These volumes of contaminated products must be sent back to a refinery to be properly separated, which significantly increases the overall operational cost of the system ([137] cited by [55] and [92]). Second, if two products are known to generate high interface losses, the PS should not place them adjacently.

The *influence of the interfaces on the energy consumption* has actually been investigated considering monthly power consumption data of the system studied [32]. Also called *the interface costs*, they correspond to the costs of reprocessing interface material between two consecutive batches [42,91]. This cost should have a value for every pair of different products [42,37,39,50] and the interface volume is also dependent on the two products [38,39,50].

Interface costs are used as part of the optimization criteria in a significant portion of the included literature [135,102,113,101,111,38,121,92,37,39,58,73,44,114,50,120,75,78,83,82,22,81,67,88,106,91,59,16,105,63,71,46,138,93,86]. According to [37,50], the product contamination at the interface of two consecutive batches of different products requires additional operations to separate them. These are expensive operations and their cost should depend on the pair of products involved in the formation of the interface. A significant set of previous studies considered the interfaces that are formed during the injection process as a natural result of the sequencing of the product, [29,51,27,55,113,111,43,117,37,50,91].

The cost of reprocessing the interfaces that are formed during the pumping operations downstream the line includes its transportation cost back to the refinery using tanker trucks and later reprocessing to separate the mixed liquids [113,38]. Also, some authors report the cost of using a *plug product*; which consists in sending a batch of intermediate product to separate two batches of non miscible products [139,99]. In either case, it significantly increases the cost of operating the system. The volume of the interface depends also on the pair of products that are involved [39]. This cost has been denominated as *transition costs* in a significant set of references [33,36,52,29,53,51,27,54,55,94,43,118,117,78,60].

Transition costs should have different values for each pair of consecutive products [38]. The amount of degraded product should be minimized but sometimes it has to be tolerated to satisfy the balance of masses [45]. According to [83], what has to be minimized is the volume of contaminated product. Nonetheless, for [46], the interface cost is independent of the interface volume, which is assumed to be constant over the course of time. In this work, however, not only the interfaces created for all touching batches at the time they are injected into the line were considered, but also the new interfaces that may be generated between old batches due to the appearance of empty batches or discharging operations downstream the line. Previous studies had focused only on the interface cost between new batches and their immediate predecessors at the time products are injected into the line.

Quite related, although not explicitly the same, quality control costs were also included in the objective function in a recent study [90]. Those are uniform unit costs for the products entering the pipeline and are associated with the measurement of chemical specifications of all the batches pumped into the pipeline during the time horizon. Furthermore, recent undertakings have also considered the product change operations as a source of costs associated to the contamination volume generated between two batches of different products [80,65].

5.1.3. Inventory costs

From the inventory optimization perspective, there are two central quantities that must be modeled. First, *the average level of inventory*. Second: *the average level of back-orders or shortages*; in case these are tolerated. In both cases, the representation scheme goes hand in hand with the behavior of demand. These two quantities must be determined for each cycle of demand.

Back-orders and shortages are not necessarily the same thing. A back-order might be a product which demand is satisfied at a point in time later than it occurs and a shortage might result in a lost sale, that is, a demand that will never be satisfied. In both cases an increase in the costs of the inventory management policies shall be expected. A few references elaborate on this kind of discussion and are commented in this section.

The minimization of the costs of keeping the products at the locations of the system where they must be stored is considered in a meaningful portion of the research work included in this review. Nonetheless, in the literature no reference was found neither addressing the demand behavior nor proposing a set of reasonable assumptions to arrive at a mathematical expression for either the average inventory, back orders or shortage levels.

Non delivered on time products have been part of the optimization criteria of a PS in a significant portion of the research included in this survey [55,102,113,107,94,111,43,117,121,62,58,73,114,50,75,103,78,83,82,67,129,106,91,59,105,63,46,86].

The minimization of the total pumped volume is directly related to both pumping and inventory costs [118]. To determine inventory costs, the mean inventory level is often used [118,42,37]. These average product inventory levels are considered at the time instants when new batches of product are pumped into the poly-duct [92]. The costs to keep inventories at the refinery and at the depots have been considered to be inversely proportional to the demand level [29] and might also be different for each commodity and location across the system [38,39]. The former can be addressed with better inventory modeling schemes with a closer consideration of the demand behavior; complemented with tailored forecasting applications for this kind of commodities.

In continuous time domains, the average inventory level of a product at a given terminal has been approximated by the sum of the product inventory levels at the end of every pumping run divided by the number of runs. Consequently, the total cost of holding inventories results from multiplying the estimated average inventory level of a product in a terminal by the cost of carrying a single unit of that product at that terminal over the horizon planning period and then adding up these costs for all products and all depots [50].

If the pumping runs have different lengths and they are not used to weigh them, then this scheme becomes a rough estimation of the inventory holding costs that can generate a progressively distorted view as subsequent schedules are generated. Moreover, it provides a discrete view of the inventories, which is a rough picture of their evolution through time when a continuous domain is adopted. In this context, it is worth noting that a recent study already weighted the costs of keeping inventories by the relative duration of each batch in the system [79].

Another scheme to estimate the inventory holding costs divides the amount in storage by the number of time slots in the grid and then multiplies it by the inventory costs [46]. Other schemes to handle inventory related costs have sought the control of inventories of every product at every center of distribution by minimizing the surplus/shortage volumes with respect to the target inventory levels [127,80,65,90]. This strategy is more a managerial approach to trace inventories than a modeling and optimization decision making scheme for inventory policy.

Inventories have also been considered from an operational perspective: some undertakings have proposed to maximize the product inventories at the end of the horizon planning period in order to deal with uncertainties [24,110,49,34,83]. This kind of approach, however, is quite contrary to what the inventory optimization paradigm recommends; which is the reduction of inventories at the levels that are just enough to handle demand at the lowest possible cost for the system.

A mathematical approach either to determine a reasonable approximation to the average inventory kept or the average tolerated back-order level of every product at any location on the system has not been introduced yet in any of the research undertakings included in this survey; neither for the deterministic, nor for the stochastic demand cases. These two mathematical models are *essential* in any realistic inventory optimization model. Moreover, what the results are from a performance evaluation of inventory review schemes (e.g.: *periodic* versus *continuous*) for this type of system remains an open question.

5.1.4. Operational criteria

The expected rate of use of the pipelines is often rolled from previous planning periods [55] and is considered within the total hours of the time horizon [121,83]. This is a result of the operational difficulty in managing the pipelines with higher utilization rates [127]. According to [125], the optimization of this criterion is accomplished by minimizing the make span. An important portion of the research work incorporates the optimization of the make span [126,22,125,57,105,18,122,104,26,65,84,47], or the amount of product pumped [24,110,116,49,118,34,128,127,48,45]. Quite interesting, the two possible directions for optimization have been considered for the latter. While for some works the maximization of this measure of effectiveness is required [24,110,116,49], other undertakings have sought its minimization [118,48]. Somehow related with this, the total pumping extent maximization has also been a key measure of effectiveness for quite a few studies [24,110,116,49,34,76].

On the other hand, the minimization of capacity under utilization [55,113,111,121,127,58,83,125], the number of cut operations [35,85,92,112,87,45,126,65], and the mean flow rate deviation have also been used as part of the optimization criteria [45,115,88,80,60,65]. In much less proportion, the minimization of storage capacity violation [127,45,126,88], the idle volume [35,85,87], the tank changeovers [130,12,23], the time window (TW) violation [20,124,126], the fixed cost of performing a multi cut operation [56] and the number of pipes reverting flow [127] have also been part of the measures of effectiveness of interest.

Split deliveries of a batch of product have been modeled in some works, also known as batch stripping operations [51,55,118] or cuts. According to [119], "*An individual cut is characterized by a single receiving depot, a unique giving batch, and the delivery size. If the pipeline is operated on fungible mode, the part of the batch to be diverted towards each depot should also be specified*". It turns out that the total number of cut operations required to meet the specified terminal demands has been an important operational optimization criterion as well [35,85,87]. This operation may have a fixed cost associated with it [112] and may also be imposed a TW to allow its occurrence. The violation of TW limits for the cuts to be performed should be minimized [126].

The operation of the pipelines at an stable flow rate is achieved by minimizing the deviation from the limits of the acceptable range of flow [45,115]. This low variability also contributes to an smoother refinery operation [88,65] and so it is a desired condition, operational wise [80]. According to [115], this state can be achieved by minimizing the magnitude of pumping rate variations in each pipeline segment along the pipeline during the scheduling horizon. It reduces the frequency of pressure variations of each pipeline segment and the number of shut downs and restarts of the pump stations along the pipelines. Consequently, the cost of operating the pipeline will also be decreased. In line with the minimization of deviation from the target limits for the flow rate, [60] proposed to use the highest pumping flow rate to maximize the amount of product transferred to the terminals.

The total pumping extent has been used as part of a weighted sum of optimization criteria in all the undertakings that have considered it as part of their measure of effectiveness. In contrast with the work of [60], the references that have considered this as an optimization criteria aim to achieve product replenishment at the lowest flow rate possible.

The TW are set to define the limits in time for the release of a batch of product from an origin and the time interval it is accepted at its destination. The definition of this parameter for the system provides the *deadline* dates for the sending and receiving operations. Violations to these TW become unavoidable in some cases as a result of the complex dynamic nature of the system, which makes the problem unfeasible if TW are strictly enforced. As a result, a soft TW scheme is proposed which aims to minimize the violations to these TW [124]. A more detailed approach defines two types of TW: *capacity TW* and *cut TW*. Capacity TW were previously defined [124] but cut TW is defined as an adjustment of the capacity TW obtained by cutting them considering the minimum time to move a batch of product [126].

According to [118], the most challenging part of the problem is the correct management of the distribution depots due to the fact that lower and upper bounds on the tank capacities must be enforced and it may take a considerable amount of pumping time to start filling them again. This is an issue of great concern that is tied to the inventory policies in use and should be a central element to include in their optimization. A lack of product in the tank affects the satisfaction of the demand of local consumers whereas an excess of it may paralyze the transfer in the pipeline and even interrupt production in the refinery. This criterion is strongly related with inventory management and its operational implications in this form of distribution system. Storage capacity should be taken into account for each product, yet it can be violated at the expense of a penalization in the objective function. This is proportional to the amount of product that exceeds it [45]. A later study defined two types of TW, derived from minimum and maximum inventory levels and capacity, as well as zero limits for inventory levels. A violation on capacity limits suffers a bigger penalty than one on the minimum/maximum inventory limits. If capacity TW is violated, inventory limits are also violated. Weights to penalize the solution are defined, according to pumping/receiving delays or earliness to origin/destination areas. The size of the weight represents the importance that is given to capacity TW violation variables. Those are weighted by the batch volume and therefore, more importance is given to violations of involving batches with bigger volumes [126].

There are two different kinds of pipelines based on their capability to handle the direction of the flow: unidirectional and bidirectional. In contrast with the former, the latter is capable of handing operations to reverse the flow of products [37]. The operation of reverting the flow inside a pipeline requires a significant operational effort and so this procedure should be minimized [127] Tank changeover costs are determined by the election of the tanks to be used to perform the pumping operations both at the refinery and at the tank farms downstream the line. It has also been considered as part of the operational optimization criteria in some undertakings [130,12,23].

5.2. Model assumptions

Before starting this discussion, it is important to remark that there are *baseline assumptions* that have been explicitly made in a significant portion of the included references related with the dynamic characteristics of the system. These are classified in two groups: hard and soft assumptions. Baseline hard assumptions define basic features inherent to the dynamics of the system operation related with the *products*, the *pipelines* and the product *contamination* that occurs inside of them. On the other hand, baseline soft assumptions are related to the system morphology, its capacity to respond to the demand of the different products and the horizon planning period. From a practical perspective, discussing these assumptions contributes to reveal the kind of response from the scientific community to provide solutions for this problem and the specific conditions of the systems to which they aim to address.

Perhaps, the two most popular and necessary hard baseline assumptions are related with the *products* and the *pipelines*. About the former, all products are assumed to have *constant densities* [52–54,114,84], which results in *liquid incompressibility* [55,5,94,111,43,35,117,121,112,87,58,44,114,56,75,103,78,74,83,72,82,69,115,91,77,21,80,61,16,60,28,64,63,84]. About the latter, pipelines on the system are assumed to remain full of *incompressible* refined products at any time during the horizon planning period [52,53,55,54,5,94,111,43,35,117,121,87,127,58,44,114,120,75,56,103,78,74,126,83,72,82,115,88,129,106,91,80,67,60,59,138,63,104,65,93,84]. Although this assumption ignores the variability of the physical properties of the fluid, as a result of the changes in the temperature and the pressure [64], it also allows to enforce the volume of product to be taken out from the line at a downstream location to be *fairly* equal to the volume of product injected at an upstream source into the line [55,111,43,37,56,75,72,68,77,21,63]. This assumption has contributed to facilitate the modeling process of the system.

The third baseline hard assumption that has explicitly been made quite often is that the batches of products inside the line are moved by the injection of new batches of products without any separation device among them [111,43,117,114,65], which results in product contamination; also known as *interfaces* or *transmix* [117,87,50,75,83,72,69,68,115,88,129,106,91,77,21,80,60,28,64,104,71,138,65,93]. Due to the high reprocessing costs of certain interfaces, some sequences of products have been considered to be forbidden [121,114,120,75,72,82,88,106,129,91,80,71,93]. Nonetheless, for some of the research work, these volumes of contaminated products have been considered to be either minimized in the constraints of the sequence [110], or not having relevance to the operation of the pipeline [64]. The latter is a contrived assumption given the high reprocessing costs of the volumes of product contaminated. Furthermore, some studies have considered the volumes of these interfaces as negligible when compared to lot volumes [110,56,74,28,64,84]. In contrast with these schemes, other works have considered these volumes to be important to decide the way in which the system should operate by assuming them to be a known constant, however independent of the pumping rate, the traveled distance and the stoppage time. Interestingly enough, the details about whether or not they depend on the pair of adjacent products have been left out of the discussion [55,111,43,117,88,106,80]. Although this kind of contribution may be of questionable relevance from a practical perspective, it is also of tremendous value from a theoretical point of view. In line with these undertakings, other works have been more explicit by clearly stating that although the volume is constant, it also depends on the pair of adjacent products [121,87,71,138,93].

There have also been some undertakings that have stated the facts about product contamination, yet the details of their corresponding assumptions about the journey of the contaminated volumes as the horizon planning period progresses are not clearly defined. [67,21,77] all state that contamination occurs between two adjacent batches of different products producing a volume of contaminated product called interface, but they do not define the behavior of the volume of these interfaces of product mixes.

In contrast with the baseline hard assumptions which have only one possibility, the baseline soft assumptions, regarding the system morphology have had quite a few different versions. A reader can come across any of these assumptions in any of the included references in this review.

The most cited references from the included research work addressed systems of different levels of complexity. These complexity spectrum ranges from single unidirectional pipeline configurations, with either a single source [51,52,55,119,112,56,75,68], or multiple sources [111], to more complex system morphology of networks of unidirectional pipelines; with either single [117] or multiple [43] sources. As previously discussed in Section 4.1, without loss of generality, the system morphology of *any instance* for this problem can be represented as a network where its nodes represent the locations of the system and its arcs represent the pipelines linking these locations. These network models can then represent different any system morphology; among which representations for systems with single pipelines, *tree like* networks of pipelines, and more complex network configurations can be found [134].

One feature of crucial modeling importance when optimizing the distribution of products is the system capacity to fulfill their demand at the consumption nodes. This capacity to provide each product can either be less than, greater than or equal to the corresponding product demand. What is more, it can also drive the measure of effectiveness adopted to make decisions on the distribution system. This is the subject of the following soft assumption.

It has been assumed through the years, in a significant portion of the research work included in this survey, that the production rate of the different products at the refinery is known [52,53,100,54,94,117,121,87,58,75,103,78,82,88,80,22,61,9,59,28,63,71,65]. Furthermore, [9] assumes that a fixed proportion of various refined products is produced. However, in the works of [80,71], an unlimited production capacity is assumed.

On the other hand, the demand rate have also been assumed to be known [52,53,100,54,94,111,43,121,58,45,75,103,78,72,82,115,88,22,77,80,61,59,9,28,63,65,84], either on a daily basis [83], constant for all products [79] or known on an hourly basis every day [87]. Only the work of [129] has explicitly stated that product demand is a *random variable* with a *uniform distribution*.

Enough system capacity to satisfy the demand requirements for each product has been assumed in the works of [130,139,63] while [61] seeks the minimization between the product demanded and the product delivered. The rest of the references did not explicitly state this relationship between supply and demand, which is a feature that a realistic application should consider.

With respect to the horizon planning period, it has been assumed to exceed the largest delivery lead time [55,91]. The length of the horizon planning period has also been a known parameter of the problem [106,91,138,80,65].

On the other hand, a feasible aggregate plan of operations for the pipelines has been assumed to be available in some works [119,112,56,74,9,28]. What is more, when simulation has been used as the main methodology, the schedule of pipeline operations has always been assumed to be already available [15,119,56]; which can be expected as simulation is not an optimization method and it can not provide a solution for the problem by itself. The aim of that kind of undertakings has been to facilitate the visualization on how the system would respond if the available schedule is implemented.

For modeling purposes, from this point forward it is convenient to classify the assumptions according to their relationship with each of the components of the system and its dynamic characteristics, which were all discussed in section 4.1.

5.2.1. Assumptions related to locations

The assumptions from this group are related with *tanks*, *injection terminals* and *receiving ports* for the most part and are discussed in that order.

There have been some basic simplifying assumptions made in many research undertakings related to tank farms. One of them has been to consider them as a system location with a capacity representing the *aggregated capacities* of all the tanks in that place. [52–54,94,121,127,50,126,83,88,22,80,59,63,65]. Some studies have assumed the set up times to switch from one tank to another in the refinery to be negligible [53,54,94,87,84]. In contrast, in the work of [68] the input station has multiple tanks to store products; as well as in the works of [9,104], where any station has multiple tanks to store products. Moreover, the state of the tanks has also been modeled by assuming it to be in one of the following possibilities: *filled to the total capacity*, *settling* or *distributing* the assigned product to clients [110,60].

Another simplifying assumption related to tanks has been to ignore their inventory related constraints [69,115,88,21,80,61,16]. On the contrary, some works have assumed upper and lower bounds for tank capacities that must be respected [139,6,45,75,22,59,28,64,71]. It has also been stated that a product must always be stored in the same tank [130,139,5,6,87,45,9]. Nonetheless, unlike with these, [45] assumes that tanks are not dedicated to a single product.

The number of tanks connected to the pipeline at the refinery to be emptied, or at all depots to be filled, has also been limited to one at any time during the horizon planning period in several works [130,139,52–54,5,94,6,87,60].

The product injection operations have also been the subject of a fairly diverse set of assumptions. In some works, for example, it has been assumed that only one *facility* injects product into the pipeline [111,119,121,120,78,21,61,28,63,138]. In line with this, other works have assumed that a single batch injection can be performed at a *supply terminal* [56,78,72,82,61,9,63,138,93]. On the other hand, the way in which the batch reception operations are performed downstream the line has also been addressed in the model assumptions made in most of the included literature. An example of this is the assumption that only one *facility* can receive product [121,87,58,120,72,129,21,61,9,28,63].

However, although these assumptions have been made in an important sub set of the included references, all of them have already been relaxed in other works. The possibility of *simultaneous product injection* at input terminals was assumed in the works of [38,43,39,103,78,74,72]. The *simultaneous product reception* at distribution centers was made possible in the works of [112,56,74,68,115,88,104,138,65] and even the possibility of *receiving multiple batches of products at the same distribution center* in the works of [80,65].

[37] investigated a system where the distribution centers have capabilities both *to receive* and *to ship* products. These types of facilities are represented as dual depots for which they proposed some assumptions to better define their operational conditions. The first assumption worth mentioning is that a depot providing products to several pipelines should also have at least one tank feeding each of them [37]. Furthermore, in this kind of facilities, pumping operations must be synchronized in the pumping runs so then incoming and outgoing flows are well coordinated. These intermediate terminals are used to transfer material between interconnected pipelines by temporarily store incoming product flows. As a result, some of these terminals could act as dual nodes; where incoming flows could be received and products with destinations located downstream the line could be injected.

One issue of central concern in the operation of this kind of terminals is the possibility of overflow. With the aim of preventing this problem, some free storage capacity was assumed to be kept to compensate when the incoming flow is greater than the outgoing flow [37]. If a terminal receives product during an individual operation, it must be taken from a single batch [56,120,75,72]. Also, it has been assumed that tanks can only receive products from the pipeline at each distribution center [84].

As the injection of a new batch of product requires that somewhere downstream the line the same amount of another product to be discharged, the *farthest distribution center first served* rule is defined. It consists in assigning the batches that are placed at the entrance of the distribution centers farther away from the refinery [71].

5.2.2. Assumptions related to commodities

The initial inventories at the refinery [111,88,22,106,59,65,84] as well as at the terminal ports have been considered to be a known parameter in a significant portion of the included literature [111,43,114,83,88,22,106,59,65]. Moreover, the initial product content inside the pipes of the system has also been assumed to be a known parameter of the system [55,111,43,37,114]. [37] relates this assumption with the execution of quality control operations on the arriving batches.

Inventory constraints at the depots have been ignored by assuming tanks to have available enough storage capacity to store the products to be injected or received anytime during the horizon planning period [67,21]. However, a discussion on how realistic the maximum amounts of inventory for every product at every location on the system according to the obtained solution was not provided.

When the batches of the same grade of product can be consolidated in a single batch and sent through the pipeline to be allocated to two or more terminals, the system is said to operate in fungible mode [55,111,35,117,128]. As a result, a product request at some distribution terminal can be satisfied by diverting material from more than one fungible batch. This can be done *dynamically* as the horizon planning period progresses [55]. As new product batches are injected, a portion of a batch flowing through the pipeline can be diverted to the assigned terminal while the remainder will continue to move towards downstream destinations, which has been called in the literature *a batch stripping operation* [55,71]. In a later study using another modeling technique, this system possibility was not considered by enforcing the flow in the pipeline section located downstream of the receiving depot to be completely stopped while making the cut operation [119].

Upper and lower bounds are enforced on the volumes of the batch injections and product deliveries to distribution centers [114,75,80,63,65]. A portion of a batch of a product can be delivered to a terminal if it is at the corresponding location and if there is space to store it in the terminal. In case there is no space, then the line must be stopped until the space issue is solved [55]. There is a minimum value for the volume of each batch to be delivered in each delivery operations [65].

The unit pumping cost varies with the product and the destination but is independent of the flow rate [55,75]. The supply rate of refined products is less than or equal to the maximum flow of the pipeline [55].

[55] introduced two assumptions to describe important aspects of the system operation in dynamic applications. In rolling schemes, the planned schedule for the current period is not changed and new requests are taken into consideration for scheduling subsequent periods. The horizon length must exceed the largest lead time, which could be somewhere between one and two weeks. This prevents batches to be injected into the system without knowing their exact destination, which is a crucial part of the decision making process.

In dynamic applications, the sequence of product shipments to be performed may vary among periods as it is optimally updated [55]. The input station injecting the refined products is located at the origin of the trunk line [117], and as previously mentioned, consecutive batches move along the pipelines with no physical barrier separating them [43]. Batch stripping operations consist in diverting part of the volume of a batch to partially satisfy the demand of the corresponding product at a distribution center and letting the rest of the batch to continue downstream the line to satisfy the demand of distribution centers located at farther points in the line [55,35,117,87,91,71]. As a result of this kind of operation, in the works of [117,71], the demand of a given product at certain distribution center can be satisfied from more than one fungible batch of product. On the other hand, the assumption that the demand of a product *must* be satisfied from only one batch has also been made in a significant set of studies [112,56,75,74,72]. However, the only way to satisfy the demand of a product from more than one batch is not through batch stripping operations only. For example, in the works of [127,126], no batch stripping operations are assumed yet the demand of a product at a distribution center can be satisfied from more than one batch of product.

The satisfaction of the demand must occur before the ending of the horizon planning period [38,43,39,114,72,82]. In some cases, that has been defined on a daily basis [121,58] and it has also been specified on an hourly basis [63].

The size of a batch of a product depends on the product itself. An important portion of papers has assumed that the size of a batch of product must be between a lower and an upper bound determined for each product [121,58,114,75,78,82,88,80,63,71,65].

5.2.3. Assumptions related to pipelines

One of the first assumptions made related to pipelines was the division of their segments in batches of given volumes [33,36,53,135,54,94,38,39,44]. Each segment of the pipeline is divided into L_d packs with equal volumetric capacity [53,54]. In the work of [100], the initial inventories inside the pipelines have not been taken into account based on the assumption that the content of each pipeline at the end of the planning interval is exactly the same as it was at the beginning. On the other hand, a significant set of works have assumed the initial inventory in the pipelines to be a known input data to build the model for the problem [111,43,114,103,83,88,22,106,129,80,59,65]. Moreover, in the work of [93] it is explicitly stated that it corresponds to the previous final state. It has been assumed that the flow rate is turbulent in order to retard mixing [55,111,91]. One popular simplifying assumption is that the flow rate is fixed during the entire time horizon [110,67,77,21], which was relaxed in the works of [56,114,120,75,72,68] where an admissible flow rate range was given for every pipeline as well as for pumping operations and receiving operations at distribution centers [65]. In the work of [56], the mean velocity was obtained from the ratio between the pump rate and the pipeline section. The electric power to drive a pump depends on the flow rate, the liquid density and the differential height [56]. The only variable determining the friction factor f at every segment is the flow rate due to the assumption of constant density and viscosity of any given oil product transported through the line [56].

The flow rate of the pipeline has to be between lower and upper bounds [112,114,120,56,75,78,74,83,72,68,88,22,129,77,16,80,59,65,86]. Product batches are sequentially pumped into the pipeline at turbulent flow to retard mixing [55,111,56] with a *Reynolds number* above 4000 [56]. In this study, the Fanning factor was used to determine the relationship between the head loss due to friction and the pump rate in a pipeline segment according to Darcy's equation [56].

Pipelines receive material from only one source. During a pumping run, a pipeline can receive material from either an adjacent pipeline or the tank farm at the pipeline origin, but not both [43]. In pipeline networks, there is at most a single pipeline between any pair of nodes in the network [43]. This assumption was relaxed in a later study where some individual pipelines can share either the source node or some destination nodes [37].

In pipeline networks, it has been assumed that all branches emerge from the mainline and so there are no branches in secondary lines [120] Programmed pipeline stoppages to run examining and repairing activities are known before the injection schedule is developed [68].

More recently made assumptions include the consideration of flow rate/pump rates at pipeline segments/refinery to be subject to given lower and upper bounds that reflect the technical limitation of the system [83,88,65,129,93,86]. However, a constant rate of flow was recently considered as well [87,62,19,28].

5.2.4. Assumptions related to flow rates

The flow rate has been considered to be fixed in a set of previous studies among the included research work [110,87,44,79,9,28]. In line with these, for [77] the flow rate must remain constant within a certain time window. If an initial station injects a batch to the pipeline, several delivery stations can simultaneously receive products from the main line. If a particular delivery was made at full flow rate, then the flow downstream the line would be stopped and as a result, the cost for flow restarts in pipeline segments would be severely increased. Consequently, full stream delivery operations must be forbidden at any intermediate delivery station, and for the same reason but with upstream consequences, injection at full stream from intermediate input stations must also be forbidden [115]. More details on the required nature of the flow rate at receiving terminals are defined in the works of [50,88], where the flow rate is determined based on the type of product at the receiving terminals. Nonetheless, it has also been assumed that the flow rate permitted by each terminal must be respected [114,68,88,22,21,80,16,59,65,93]. Furthermore, in the works of [69,67,61,28], the flow rate at the terminals has been assumed to be constant. On the other hand, the rate of flow at the terminals has been assumed to be properly determined through the regulation of the operation of the pump within a time window [77]. What is more, more details have been recently provided by assuming that the maximum delivery rate to a certain center of distribution may be lower than the maximum flow rate in the immediate upstream segment, as a result of the impossibility of some products to be delivered at full pressure [86].

On the other hand, the source has also been considered to determine the flow rate as not all sources have the same injection capacity to dispatch products downstream the line. This flow rate has been required to be within a lower and an upper bound according to the shipping terminal in a significant set of works [111,43,117,72,80,59,22]. However, for the works of [38,82], the injection of products must be performed at a constant rate of flow, while in the work of [16], it has to be constant within a single time window.

The flow rate has also been determined based on the product that is being either injected or received. Each product has a lower and an upper bound for the flow rate that can be used to move them through the lines of the system and these limits must be respected in the programming of the transportation operations [102,5,88,59]. However, recently the works of [79,9] have assumed a product independent constant flow rate.

A linear relationship between the flow rate and the required energy for the operation of the pumps has been assumed [61], which is in line with the constant nature of electricity prices assumed in the work of [64].

6. Emerging and future research trends

Pipeline sharing applications have been emerging in China as a topic of great interest, perhaps as a result of the convergence of the need of a more efficient use of this kind of expensive infrastructure, and the demand for the use of alternative and more sustainable sources of energy; which can be transported in this type of infrastructure. Recently, [31] provided a framework for the optimal remaining capacity estimation of a pipeline system optimizing the revenue generated from selling that capacity for the transportation of other products. Another study on the pipeline sharing topic developed a logistics optimization model considering multi product pipeline scheduling from the perspective of biofuels suppliers [7]. This topic aims to provide decision making tools for the use of already existing pipeline infrastructure in the transportation of biofuels. More recently, another framework for this purpose was published in the context of the Chinese fuel transportation logistics. Oil depletion and the demanded transition to a low carbon economy are encouraging the consideration of biofuels as effective alternatives to fuel oils and so multi product pipelines will be required to broad its portfolio of products transported by including biofuels [131]. But not only biofuels must be considered as possible products to be transported in a pipeline system. The use of multi product pipeline systems for the transportation of other substances, such as green amonia, a promising energy storage form, is already being investigated [8].

The included research work identifies research avenues on two central paradigms for the problem: modeling and solution procedures. This section is devoted to present a discussion on the research avenues identified by the authors of the included references in this survey in these two lines of study. First, it provides a discussion on the future modeling challenges proposed by the authors of the included references. Second, it remarks the challenges posed by the solution procedures, as identified in the included research work. Third, considering the need of better insights into the decision making process for this economically important problem, it also proposes what the fertile future research avenues are that should receive an increasing amount of attention in the coming years.

6.1. Future work in the modeling paradigm

The three most frequently mentioned research avenues to contribute to the modeling paradigm are: to extend the proposed model to a *more complex morphology*, to include *more complex operational aspects* for the model and to have an *enhanced version of the objective function*.

In the case of future model extensions to address a more complex morphology, although some works that expressed their interest in this line of work did not explicitly state the details of their future research endeavors [102,87,58,120,72,88], a significant amount of studies centered their future attention in systems with a tree shaped network of pipelines [43,117,75,74,68,57,63]. Moreover, the so called mesh structure network of pipelines, has also been identified as the center of future implementations [50,129,26]. Those system configurations are both applications of *network morphology* on the system and their study seems to be of interest in other works, yet without providing much level of detail [132,82,140,105,122,104].

An examination of the research work included in this survey reveals that some of the most important operational aspects have consistently been left out from the models proposed for the system in a significant portion of the included references. Although their inclusion has been beyond their scope at the time they were published, considering these key aspects is due in the coming research undertakings. One example of this kind of research avenue is the study of the effect of the different sources of randomness on the system. Among the most popular feature in this type of applications is the consideration of the stochastic version for the demand of the products, which is perhaps one of the most important lines of work in this research avenue. A considerable number of studies have stated as their future research avenue the modeling of this kind of behavior in the product consumption by the customers of the distribution system [101,42,121,37,83,81,106,61,93]. In line with this, considering the time at which the demand of each customer occurs has also been mentioned as a topic of future interest in research [21]. Other sources of randomness which effects might result in unforeseen pipeline stoppages that would affect the normal operation of the system have also been suggested [44,76,61,63]; however the details on the nature of these randomness sources and the effects they generate on the system have not been discussed.

Another complicating operational feature in some instances of this kind of distribution system is the capability of certain pipes to perform the so called *reverse flow operations*. Future research efforts to investigate on how to handle this challenging system feature in new contributions have also been mentioned in a significant portion of works from the included literature [102,43,44,50,57,79,105]. Although the inclusion of this kind of operational feature would apparently allow the consideration of a broader set of dispatching options; perhaps resulting in the ability to find even lower cost schedules, it might also result in a double number of decision variables in the model [44]. This would contribute to create a larger search space for the optimization solver. Far from being a trivial exercise, it configures a challenging line of work for future research efforts, not only from the modeling perspective, but also for the solution procedures realm.

The modeling of incoming and outgoing flows at transfer terminals [103] as well as at the refineries [84] have also been proposed as a central topic in future research works. They are in the same direction with the interest in including the modeling of simultaneous injection and delivery operations downstream the pipelines [103]. Slightly modified, other studies have also identified in their future research efforts the inclusion of multiple sources [10,114,120,79,122,65] and distribution centers [83,81,47]. In the same line of work, another future research effort will take care of the generalization of their modeling scheme to a more complex system where a given pipeline can receive product not only from a source terminal but also from the previous pipeline; both occurring during the same period of time [72]. Closely related, another work centers their future attention in the study of the performance of their proposed model when allowing multiple batches of products to be simultaneously transferred to two branches of a segment [59].

Modifications on the objective functions have also been mentioned as a topic of interest in future research efforts. The use of different objective functions is proposed to be investigated by [118]. Perhaps in line with this, an improvement of objective functions has also been stated as future developments in a significant amount of studies. While some works have just provided a general idea without explicitly stating their aims [10,141,53,54,109,111,44,75,93], some other undertakings have stated their interest to be centered in using costs based objective functions [127,128] to guide the search process of their heuristic schemes [49,6].

The objective function elements that have been the subject of future research directions in a significant portion of the references included in this survey, which are mentioned the most, are the costs of *pumping products into the line* and *product contamination*. One example of this kind of discussion is the inclusion of electricity costs and the costs associated with product contamination [76]. In a realistic scenario, pumping costs are not constant through the horizon planning period as they are actually time dependent and that condition should be represented in the models [54]. Peak hour pumping costs have been consistently mentioned in an important set of works as a topic of future interest [10,53,54,109,111,6,44,88,65]. In a very similar line of work, other studies have suggested the use of new constraints in the proposed models to address the variability in the pumping costs [6,44]. Also, a linear approximation of the objective function is proposed as a topic included in future works [49]. Nonetheless, depending on the resulting model structure, adding more constraints might also result in the requirement of a more challenging optimization procedure. Furthermore, according to [127], it is hard to implement economic features in the objective function due to the existing lack of information from real situations.

With an essentially identical aim of including the study of pumping costs, a previous work identified as a future research avenue their optimization using discrete event systems simulation [111]. However, there are two important concerns with this initiative that are worth mentioning: first, discrete event system simulation is not an optimization technique by itself. Second, the details in its role on the optimization algorithm were not clearly discussed.

The inclusion of the optimization of the schedule at the refinery as part of the model is proposed as a future research direction [129,106,60]. Pipeline hydraulic calculations [69] as well as fluid dynamics considerations have also been stated as a promising future lines of work [65].

Other operational aspects that have also been proposed to be included in coming undertakings are related with the forbidden product sequences. While a previous study suggests modeling the use of plug products to allow forbidden sequences via new constraints for the model [80], another study identified the development of new constraints for their model to enforce the forbidden product sequences [18] as a new feature to be included in later works. Either case, the new version of the model is very likely to be more computationally demanding to solve.

The inclusion of settling periods to preserve the quality of products has also been identified as a future feature to be included in the models proposed [80,79,65].

Batch optimization consists in defining the sequence in which the batches of product must be injected into the line, and the *size* of each batch [16,62]. An important portion of the research work uses a batch volume of fixed size. However, variable sizes provide higher flexibility to the solution using less computational time [76]. To determine these two key factors translates into the optimization of the input plan at the initial station on the line [115,61]. Other authors mention this as an automatic selection of the initial batch sequence [26].

The management of the operations at the so called tank farms, where the products are either received at the consumption nodes, or sent from the dispatching depots, has also been identified as a topic that should receive increasing attention in the future. At the refinery, for example, the inventory policies at the tank facilities can seriously affect the predefined delivery sequence before the batches of product are injected into the line. A logic based thought process arrives at the conclusion that the actual order of product injection into the line depends directly on the available volume of each product that form the sequence to be pumped into the line. In most of the included literature, it is assumed that there is an enough amount of each product at the origin to be dispatched through the line. A more realistic application should then include the definition of policies to manage the inventories of different products, not only at the customer facilities, but also at the origin from which the products are dispatched. As a result, the definition of *inventory management policies* at the refinery storage tank facilities as long as at these from the customers connected downstream the lines, shall not be untied from the definition of the optimal distribution schedule in future research undertakings on this topical area.

Regarding the management of the tank facilities across the distribution system, there have been three main topics left for future discussions: first, the implementation of inventory management constraints at all the tank facilities across the system [24,110,116,6,76,21,84]. Second, the dis-aggregation of tanks at the depots [116,76]. Third, the inclusion of specific requirements of tank facilities over time. These requirements include the *settling periods* at the tanks according to the requirements of the different products [80,79,65,84], and the *availability and limits of their capacity* [49,125], which can be affected by scheduled stops to perform maintenance operations on the tanks [48]. Although, not specifically mentioning the tank facilities, other authors have also expressed their interest in addressing the modeling of inventory management constraints in their future research efforts [109].

In virtually every case that has been mentioned so far in this discussion, it is very likely that all these new features might result in *more complex versions* of their proposed mathematical models as implementing them will necessarily imply a larger set of new decision variables and constraints. This will result in a even more complex computationally expensive optimization procedure, which by itself shall become a challenging line or work inside this topical area. That is the topic of discussion in the following subsection.

6.2. Solution procedure challenges

Far from being a short crevice in this topical area, the development of solution procedures poses a fertile research avenue yet to be fully explored and developed. Future studies should point to avenues for obtaining better quality solutions in acceptable computational time for larger pipeline networks [44]; not only as the result of more comprehensive modeling schemes, but also more efficient and robust solution procedures. The solution procedures that have been investigated are predominantly optimization schemes based on heuristic and meta heuristic solution strategies. On the other hand the paradigm of *mathematical decomposition based* solution procedures [142,143] for the proposed models continues to be virtually unexplored. There has been only one undertaking in this direction [63], recently made, which reported quite encouraging results from their implementation. Apart from this work, the vast majority of the proposed mathematical programming models has been addressed as a *monolithic* solution approach by using the default settings of the solvers used while discussions on their algebraic structure and the possibilities that they offer to be exploited in a tailored mathematical decomposition based solution procedure have not been made in any case. The decomposition based solution schemes that have been used are rather centered in solving the problem by dividing the decision making process in various components that are not connected with one another in some form of exact mathematical procedures to communicate with every other component during the optimization process. Instead, heuristic rules have been proposed to guide the process leaving enough room for research efforts in the line of exact solution procedures. Benders decomposition and Lagrangian relaxation are two promising lines of work. Those solution schemes should be investigated given the large scale nature of the proposed models and the complexity of the system under consideration.

6.3. Future research avenues

Inventory management involves both the demand forecasting and the modeling and optimization of product inventories to arrive at the policies that allow their optimization in the long run. None of these lines of work has been addressed so far from an operations management perspective. Refined products inventory modeling and optimization, at the refinery and at the depots, pose a necessary modeling challenge as this representation must be part of the abstraction made of this kind of distribution system formed with strongly interconnected parts. Mathematical expressions for the average inventories, as well as average shortages, are usually strongly dependent on the demand behavior and are an essential part of most inventory optimization models. What is more, periodic versus continuous review policies have not been studied yet for neither modeling nor solution procedures purposes. As a result, the inclusion of inventory modeling capabilities requires a close consideration of product demand behavior at every location of the system. All the research work that was included in this survey has been driven by a push scheme to operate the system. Inventory management policies have always been a side effect of the pipelines schedule but an scheme for the opposite order on the decision making process has not been investigated yet neither in the modeling nor in the solution procedure research avenues. A pull scheme for the system operation poses a promising line of work as it remains unclear whether it could arrive at a better system performance or not. This kind of operational design should first determine the product inventory policies at all locations of the system and then the PS to fulfill those policies. Under this scheme, the settling periods and transportation time should be considered to determine the products lead time.

In spite of the inherent network structure of this system, a successful representation of it with a network flow model is still due in the literature of this topical area. None of the included references in this review proposed this kind of undertaking, perhaps because the nature of the operation of the pipelines results in a set of dynamic properties that complicate their modeling using a scheme of representation based on a plain network flow model. A challenging set of yet to be proposed additional constraints on the flows through the arcs must be proposed to model these unique features only found in this kind of distribution system.

Although the research effort has increasingly gotten closer to more realistic modeling paradigms for the system, as discussed in section 5.2, the research contributions in the optimization procedures continue to be particularly scarce, as far as *exact mathematical optimization procedures* are concerned. Research on the implementation of meta heuristic solution schemes is also limited. This line or work should become a required fertile research avenue for this problem during the following years.

7. Concluding remarks

Up to date, a prominent portion of the attention from researchers has been received by *single unidirectional pipeline systems*. The top three sources of publication of this kind of work are the *Computers & Industrial Engineering* journal, the *Industrial & Chemistry Engineering*

Research journal and the IEOR-AM domain. The top five most cited journals from the latter are *Annals of Operations Research, Applied Mathematical Modeling, Computers & Operations Research, European Journal of Operational Research, Omega, Computers & Industrial Engineering* and *Optimization & Engineering*. This domain has been becoming a significant contributor in both papers and citations (Fig. 4). The number of publications addressing this kind of configuration, as well as their relevance, is bigger by far when compared to their network of pipelines counterpart.

The transportation of refined products using pipelines presents two significant research challenges that need to be addressed to optimize this system: *developing robust mathematical models* and *creating effective solution procedures*. The chemical nature of the products being transported, the system dynamics complexity and the high computational effort to solve the models without exploiting their algebraic structure on a tailored solution procedure demand an interdisciplinary approach to address this problem. The solution of this challenge stands on the interface of CE and OR. The former must guide the system modeling component while the latter must provide efficient solution schemes for the optimization of proposed models.

The modeling paradigm should involve two additional elements from the operations management field: forecasting and product inventory systems. The consideration of these two features adds an additional layer of realism to the modeling paradigm of this kind of transportation system; and of complexity as well. So far, the research work that has been developed in this area has not considered this two features of great importance in the proposed models. In static applications, a discrete view of product demand behavior is used as input data. In dynamic applications, product demands are being updated with the course of the horizon planning period but the use of forecasting methods has not being reported to come up with these data. Demand behavior is strongly connected to average inventories and so its study will facilitate a better modeling of product inventories; for which rough approximations have been made.

Despite some relatively simple morphology have been investigated in the included literature, this kind of transportation system might also be found in larger sizes. For example, the *Colonial Pipeline* system spans over 5500 miles from Houston to New York and has about 65 stub lines to serve different regions in eastern United States. It transports a wide variety of refined products with seasonal demands across the region it serves. This system is much bigger than any of the instances addressed in the included literature. What is more, a meaningful portion of the research work states that scaling up their models would make the required computational time to find a solution to become prohibitive. Perhaps the best approach to solve those models might not come from a single type of scheme but from a combination of more than one optimization paradigm. This might involve the use of meta heuristic solution procedures combined with mathematical decomposition schemes.

Although the CE sphere continues to be the predominant domain of the publications on this topical area by far, preserving the vast majority of the citations as well, the IEOR-AM realm has been gaining a meaningful participation and relevance on it. The number of publications from this domain, as well as the number of citations of these publications made by the relevant literature, have been substantially growing with the course of time. From the supply chain point of view, this problem has a vast economic importance and poses a progressively purposeful challenge for the IEOR-AM domain for undertakings centered in both modeling and solution procedure research efforts.

CRediT authorship contribution statement

Rolando José Acosta-Amado: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data associated with the included papers are available in the original sources. Full supplemental data from the review process will be made available to the interested readers upon request.

References

[1] T. Denton, Pipeline safety excellence performance 2019 annual liquids report, Tech. Rep., Association of Oil Pipelines AOPL, American Petroleum Institute API, 2018.

[2] S.A. Yatauro, 2023-2025 pipeline excellence strategic plan & 2022 performance report, Tech. Rep., The American Petroleum Institute, the Liquid Energy Pipeline Association - LEPA, 2023.

[3] J. Stoody, Us liquids pipeline usage and mileage report, Tech. Rep., Association of Oil Pipelines AOPL, American Petroleum Institute API, 2014.

[4] M. French, N. Ameen, U.s. crude oil production forecast to rise in 2022 and 2023 to record-high levels, https://www.eia.gov/todayinenergy/detail.php?id=51318#, February 2022.

[5] A.V. Moura, C.C. de Souza, A.A. Cire, T.M. Lopes, Planning and scheduling the operation of a very large oil pipeline network, in: Principles and Practice of Constraint Programming, Springer, 2008, pp. 36–51.

[6] T.M.T. Lopes, A.A. Ciré, C.C. de Souza, A.V. Moura, A hybrid model for a multiproduct pipeline planning and scheduling problem, Constraints 15 (2) (2010) 151–189.

[7] R. Tu, Q. Liao, N. Xu, X. Wei, Y. Wang, Y. Liang, H. Zhang, Pipeline sharing: potential capacity analysis of biofuel transportation through existing pipelines, J. Clean. Prod. 398 (2023) 136507.
[8] R. Tu, C. Liu, Q. Shao, Q. Liao, R. Qiu, Y. Liang, Pipeline sharing: optimal design of downstream green ammonia supply systems integrating with multi-product pipelines, Renew. Energy 223 (2024) 120024.
[9] A. Siddiqui, M. Verma, V. Verter, An integrated framework for inventory management and transportation of refined petroleum products: pipeline or marine?, Appl. Math. Model. 55 (2018) 224–247.
[10] C.A. Hane, H.D. Ratliff, Sequencing inputs to multi-commodity pipelines, Ann. Oper. Res. 57 (1) (1995) 73–101.
[11] L. Magatão, L. Arruda, F. Neves, A mixed integer programming approach for scheduling commodities in a pipeline, Comput.-Aided Chem. Eng. 10 (2002) 715–720.
[12] L. Magatão, L.V. Arruda, F. Neves, A mixed integer programming approach for scheduling commodities in a pipeline, Comput. Chem. Eng. 28 (1) (2004) 171–185.
[13] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: a software environment for integrated models of biomolecular interaction networks, Genome Res. 13 (11) (2003) 2498–2504.
[14] F. Neves-Jr, L. Magatão, S.L. Stebel, S.N. Neves, L.C. Felizari, D.I. Czaikowski, R. Rocha, P.C. Ribas, An efficient approach to the operational scheduling of a real-world pipeline network, Comput.-Aided Chem. Eng. 24 (2007) 697–702.
[15] F.M. Mori, R. Lüders, L.V.R. de Arruda, L. Yamamoto, M.V. Bonacin, H.L. Polli, M.C. Aires, L.F. de Jesus Bernardo, Simulating the operational scheduling of a realworld pipeline network, in: European Symposium Computer Aided Process Engineering ESCAPE-15, vol. 24, 2007, pp. 691–696.
[16] Q. Liao, H. Zhang, Y. Wang, W. Zhang, Y. Liang, Heuristic method for detailed scheduling of branched multiproduct pipeline networks, Chem. Eng. Res. Des. 140 (2018) 82–101.
[17] S.N. Boschetto, L.C. Felizari, L. Yamamoto, L. Magatão, S.L. Stebel, F. Neves-Jr, L.V.R. de Arruda, R. Lüders, P.C. Ribas, L.F. de Jesus Bernardo, An integrated framework for operational scheduling of a real-world pipeline network, Comput.-Aided Chem. Eng. 25 (2008) 259–264.
[18] P.M. Castro, H. Mostafaei, New continuous-time scheduling formulation for multilevel treelike pipeline systems, in: Computer Aided Chemical Engineering, vol. 43, Elsevier, 2018, pp. 973–978.
[19] V. Cafaro, D.C. Cafaro, C.A. Mendez, J. Cerda, Detailed scheduling of oil products pipelines with parallel batch inputs at intermediate sources, Chem. Eng. 32 (2013).
[20] L.C. Felizari, L.V. de Arruda, R. Lüders, S.L. Stebel, Sequencing batches in a real-world pipeline network using constraint programming, Comput.-Aided Chem. Eng. 27 (2009) 303–308.
[21] Q. Liao, Y. Liang, N. Xu, H. Zhang, J. Wang, X. Zhou, An milp approach for detailed scheduling of multi-product pipeline in pressure control mode, Chem. Eng. Res. Des. 136 (2018) 620–637.
[22] H. Mostafaei, P.M. Castro, Continuous-time scheduling formulation for straight pipelines, AIChE J. 63 (6) (2017) 1923–1936.
[23] L. Magatao, L. Arruda, F. Neves-Jr, Using clp and milp for scheduling commodities in a pipeline, in: European Symposium Computer Aided Process Engineering ESCAPE-15, vol. 20, 2005, p. 1027.
[24] S. Relvas, A.P.F. Barbosa-Póvoa, H.A. Matos, J. Fialho, A.S. Pinheiro, Pipeline scheduling and distribution centre management—a real-world scenario at clc, Comput.-Aided Chem. Eng. 21 (2006) 2135–2140.
[25] N. Xu, Q. Liao, Z. Li, Y. Liang, R. Qiu, H. Zhang, An efficient decomposition approach for the low-energy scheduling of a straight multiproduct pipeline, Chem. Eng. Res. Des. 165 (2021) 341–360.
[26] Q. Liao, P.M. Castro, Y. Liang, H. Zhang, Batch-centric model for scheduling straight multisource pipelines, AIChE J. 65 (10) (2019) e16712.
[27] D.C. Cafaro, J. Cerdá, Multiperiod planning of multiproduct pipelines, Comput.-Aided Chem. Eng. 20 (2005) 1453–1458.
[28] Q. Liao, H. Zhang, T. Xia, Q. Chen, Z. Li, Y. Liang, A data-driven method for pipeline scheduling optimization, Chem. Eng. Res. Des. 144 (2019) 79–94.
[29] D.C. Cafaro, J. Cerdá, A continuous-time approach to multiproduct pipeline scheduling, in: European Symposium on Computer Aided Process Engineering-13: 36th European Symposium of the Working Party on Computer Aided Process Engineering, vol. 14, Elsevier, 2003, p. 65.
[30] S. Relvas, A.P. Barbosa-Póvoa, H.A. Matos, Oil products distribution systems: decomposition approach on pipeline and inventory scheduling, in: Computer Aided Chemical Engineering, vol. 27, Elsevier, 2009, pp. 1971–1976.
[31] R. Tu, Q. Liao, L. Huang, Y. Jiao, X. Wei, Y. Liang, Pipeline sharing: remaining capacity estimation of multiproduct pipelines, Chem. Eng. Res. Des. 191 (2023) 338–352.
[32] S. Relvas, A.P.F. Barbosa-Póvoa, H.A. Matos, Optimal operation of a real multiproduct pipeline and storage system: economical versus operational objectives, Comput.-Aided Chem. Eng. 26 (2009) 399–404.
[33] R. Rejowski, J. Pinto, An milp formulation for the scheduling of multiproduct pipeline systems, Braz. J. Chem. Eng. 19 (4) (2002) 467–474.
[34] S.N. Boschetto, S. Relvas, A.P.F. Barbosa-Póvoa, F. Neves, A decomposition approach for the operational scheduling of a multiproduct pipeline, Comput.-Aided Chem. Eng. 28 (2010) 1207–1212.
[35] M. Gleizes, G. Herrero, D. Cafaro, C. Méndez, J. Cerdá, A discrete-event simulation tool for efficient operation of refined products pipelines, in: Proceedings of the 20th European Symposium on Computer Aided Process Engineering, ESCAPE 20, 2010, pp. 1–6.
[36] R. Rejowski, J. Pinto, Efficient milp formulations for multiproduct pipeline scheduling, Comput.-Aided Chem. Eng. 15 (2003) 1002–1007.
[37] D.C. Cafaro, J. Cerdá, Rigorous scheduling of mesh-structure refined petroleum pipeline networks, Comput. Chem. Eng. 38 (2012) 185–203.
[38] A. Herrán, J. De la Cruz, B. De Andrés, A mathematical model for planning transportation of multiple petroleum products in a multi-pipeline system, Comput. Chem. Eng. 34 (3) (2010) 401–413.
[39] A. Herrán, J. de la Cruz, B. de Andrés, Global search metaheuristics for planning transportation of multiple petroleum products in a multi-pipeline system, Comput. Chem. Eng. 37 (2012) 248–261.
[40] B. Csontos, L. Halász, I. Heckl, Event-driven simulation method for fuel transport in a mesh-like pipeline network, Comput. Chem. Eng. 157 (2022) 107611.
[41] R.F. Banaszewski, L.V. Arruda, J.M. Simão, C.A. Tacla, A.P. Barbosa-Póvoa, S. Relvas, An application of a multi-agent auction-based protocol to the tactical planning of oil product transport in the Brazilian multimodal network, Comput. Chem. Eng. 59 (2013) 17–32.
[42] S. MirHassani, H.F. Jahromi, Scheduling multi-product tree-structure pipelines, Comput. Chem. Eng. 35 (1) (2011) 165–176.
[43] D.C. Cafaro, J. Cerdá, Operational scheduling of refined products pipeline networks with simultaneous batch injections, Comput. Chem. Eng. 34 (10) (2010) 1687–1704.
[44] E.M. de Souza Filho, L. Bahiense, V.J.M. Ferreira Filho, Scheduling a multi-product pipeline network, Comput. Chem. Eng. 53 (2013) 55–69.
[45] J.A. Fabro, S.L. Stebel, D. Rossato, H.L. Polli, L.V. Arruda, F. Neves, P.C. Ribas, A.P.F. Barbosa-Póvoa, S. Relvas, A milp (mixed integer linear programming) decomposition solution to the scheduling of heavy oil derivatives in a real-world pipeline, Comput. Chem. Eng. 66 (2014) 124–138.
[46] Q. Liao, P.M. Castro, Y. Liang, H. Zhang, New batch-centric model for detailed scheduling and inventory management of mesh pipeline networks, Comput. Chem. Eng. 130 (2019) 106568.
[47] Y. Yan, P.M. Castro, Q. Liao, Y. Liang, An effective decomposition algorithm for scheduling branched multiproduct pipelines, Comput. Chem. Eng. (2021) 107494.
[48] T.M. Lopes, A.V. Moura, C.C. de Souza, A.A. Cire, Planning the operation of a large real-world oil pipeline, Comput. Chem. Eng. 46 (2012) 17–28.
[49] S. Relvas, A.P.F. Barbosa-Póvoa, H.A. Matos, Heuristic batch sequencing on a multiproduct oil distribution system, Comput. Chem. Eng. 33 (3) (2009) 712–730.
[50] D.C. Cafaro, J. Cerdá, Rigorous formulation for the scheduling of reversible-flow multiproduct pipelines, Comput. Chem. Eng. 61 (2014) 59–76.

[51] D.C. Cafaro, J. Cerdá, Optimal scheduling of multiproduct pipeline systems using a non-discrete milp formulation, Comput. Chem. Eng. 28 (10) (2004) 2053–2068.
[52] R. Rejowski, J.M. Pinto, Scheduling of a multiproduct pipeline system, Comput. Chem. Eng. 27 (8) (2003) 1229–1246.
[53] R. Rejowski, J.M. Pinto, Efficient milp formulations and valid cuts for multiproduct pipeline scheduling, Comput. Chem. Eng. 28 (8) (2004) 1511–1528.
[54] R. Rejowski, J.M. Pinto, A novel continuous time representation for the scheduling of pipeline systems with pumping yield rate constraints, Comput. Chem. Eng. 32 (4) (2008) 1042–1066.
[55] D.C. Cafaro, J. Cerdá, Dynamic scheduling of multiproduct pipelines with multiple delivery due dates, Comput. Chem. Eng. 32 (4) (2008) 728–753.
[56] V.G. Cafaro, D.C. Cafaro, C.A. Méndez, J. Cerdá, Minlp model for the detailed scheduling of refined products pipelines with flow rate dependent pumping costs, Comput. Chem. Eng. 72 (2015) 210–221.
[57] P.M. Castro, H. Mostafaei, Product-centric continuous-time formulation for pipeline scheduling, Comput. Chem. Eng. 104 (2017) 283–295.
[58] S. MirHassani, N. BeheshtiAsl, A heuristic batch sequencing for multiproduct pipelines, Comput. Chem. Eng. 56 (2013) 58–67.
[59] P.M. Castro, H. Mostafaei, Batch-centric scheduling formulation for treelike pipeline systems with forbidden product sequences, Comput. Chem. Eng. 122 (2019) 2–18.
[60] D. Dimas, V.V. Murata, S.M. Neiro, S. Relvas, A.P. Barbosa-Póvoa, Multiproduct pipeline scheduling integrating for inbound and outbound inventory management, Comput. Chem. Eng. 115 (2018) 377–396.
[61] Q. Liao, H. Zhang, N. Xu, Y. Liang, J. Wang, A milp model based on flowrate database for detailed scheduling of a multi-product pipeline with multiple pump stations, Comput. Chem. Eng. 117 (2018) 63–81.
[62] L. Yongtu, L. Ming, Z. Ni, A study on optimizing delivering scheduling for a multiproduct pipeline, Comput. Chem. Eng. 44 (2012) 127–140.
[63] N.B. Asl, et al., Benders decomposition with integer sub-problem applied to pipeline scheduling problem under flow rate uncertainty, Comput. Chem. Eng. 123 (2019) 222–235.
[64] X. Zhou, H. Zhang, R. Qiu, Y. Liang, G. Wu, C. Xiang, X. Yan, A hybrid time milp model for the pump scheduling of multi-product pipelines based on the rigorous description of the pipeline hydraulic loss changes, Comput. Chem. Eng. 121 (2019) 174–199.
[65] W.H.T. Meira, L. Magatão, F. Neves-Jr, L.V. Arruda, J.P. Vaqueiro, S. Relvas, A.P. Barbosa-Póvoa, Scheduling of a single-source multiproduct pipeline system by a matheuristic approach: combining simulated annealing and milp, Comput. Chem. Eng. 136 (2020) 106784.
[66] L. Yu, S. Wang, Q. Xu, Optimal scheduling for simultaneous refinery manufacturing and multi oil-product pipeline distribution, Comput. Chem. Eng. 157 (2022) 107613.
[67] H. Zhang, Y. Liang, Q. Liao, M. Wu, X. Yan, A hybrid computational approach for detailed scheduling of products in a pipeline with multiple pump stations, Energy 119 (2017) 612–628.
[68] H. Chen, L. Zuo, C. Wu, L. Wang, F. Diao, J. Chen, Y. Huang, Optimizing detailed schedules of a multiproduct pipeline by a monolithic milp formulation, J. Pet. Sci. Eng. 159 (2017) 148–163.
[69] H.-R. Zhang, Y.-T. Liang, Q. Xiao, M.-Y. Wu, Q. Shao, Supply-based optimal scheduling of oil product pipelines, Pet. Sci. 13 (2) (2016) 355–367.
[70] Y. Liang, M. Li, J. Li, Hydraulic model optimization of a multi-product pipeline, Pet. Sci. 9 (4) (2012) 521–526.
[71] S. Moradi, S. MirHassani, F. Hooshmand, Efficient decomposition-based algorithm to solve long-term pipeline scheduling problem, Pet. Sci. 16 (5) (2019) 1159–1175.
[72] H. Mostafaei, P.M. Castro, A. Ghaffari-Hadigheh, Short-term scheduling of multiple source pipelines with simultaneous injections and deliveries, Comput. Oper. Res. 73 (2016) 27–42.
[73] S. MirHassani, M. Abbasi, S. Moradi, Operational scheduling of refined product pipeline with dual purpose depots, Appl. Math. Model. 37 (8) (2013) 5723–5742.
[74] V.G. Cafaro, D.C. Cafaro, C.A. Méndez, J. Cerdá, Optimization model for the detailed scheduling of multi-source pipelines, Comput. Ind. Eng. 88 (2015) 395–409.
[75] A. Ghaffari-Hadigheh, H. Mostafaei, On the scheduling of real world multiproduct pipelines with simultaneous delivery, Optim. Eng. 16 (3) (2015) 571–604.
[76] S. Relvas, S.N.B. Magatão, A.P.F. Barbosa-Póvoa, F. Neves, Integrated scheduling and inventory management of an oil products distribution system, Omega 41 (6) (2013) 955–968.
[77] Z. Haoran, L. Yongtu, L. Qi, S. Yun, Y. Xiaohan, A self-learning approach for optimal detailed scheduling of multi-product pipeline, J. Comput. Appl. Math. 327 (2018) 41–63.
[78] H. Mostafaei, Y. Alipouri, J. Shokri, A mixed-integer linear programming for scheduling a multi-product pipeline with dual-purpose terminals, Comput. Appl. Math. 34 (3) (2015) 979–1007.
[79] T. Kirschstein, Planning of multi-product pipelines by economic lot scheduling models, Eur. J. Oper. Res. 264 (1) (2018) 327–339.
[80] W.H.T. Meira, L. Magatão, S. Relvas, A.P. Barbosa-Póvoa, F. Neves Jr, L.V. Arruda, A matheuristic decomposition approach for the scheduling of a single-source and multiple destinations pipeline system, Eur. J. Oper. Res. 268 (2) (2018) 665–687.
[81] S. Moradi, S. MirHassani, Robust scheduling for multi-product pipelines under demand uncertainty, Int. J. Adv. Manuf. Technol. 87 (9–12) (2016) 2541–2549.
[82] A. Zaghian, H. Mostafaei, An milp model for scheduling the operation of a refined petroleum products distribution system, Oper. Res. 16 (3) (2016) 513–542.
[83] S. Moradi, S. MirHassani, Transportation planning for petroleum products and integrated inventory management, Appl. Math. Model. 39 (23) (2015) 7630–7642.
[84] H. Mostafaei, P.M. Castro, S. Relvas, I. Harjunkoski, A holistic milp model for scheduling and inventory management of a multiproduct oil distribution system, Omega 98 (2021) 102110.
[85] V.G. Cafaro, D.C. Cafaro, C.A. Méndez, J. Cerdá, Oil-derivatives pipeline logistics using discrete-event simulation, in: Proceedings of the Winter Simulation Conference, Winter Simulation Conference, 2010, pp. 2101–2113.
[86] H. Mostafaei, P.M. Castro, F. Oliveira, I. Harjunkoski, Efficient formulation for transportation scheduling of single refinery multiproduct pipelines, Eur. J. Oper. Res. 293 (2) (2021) 731–747.
[87] M.F. Gleizes, G. Herrero, D.C. Cafaro, C.A. Méndez, J. Cerdá, Managing distribution in refined products pipelines using discrete-event simulation, Int. J. Inf. Syst. Supply Chain Manag. 5 (1) (2012) 58–79.
[88] W.H.T. Meira, L. Magatão, S. Relvas, A.P.F.D.B. Póvoa, F.N. Junior, A decomposition approach for the long-term scheduling of a single-source multiproduct pipeline network, in: Congress of APDIO, the Portuguese Operational Research Society, Springer, 2017, pp. 235–248.
[89] V.G. Cafaro, D.C. Cafaro, J. Cerdá, Improving the mathematical formulation for the detailed scheduling of refined products pipelines by accounting for flow rate dependent pumping costs, Iberoam. J. Ind. Eng. 5 (10) (2013) 115–128.
[90] M. Quinteros, M. Guignard, A. Weintraub, M. Llambias, C. Tapia, Optimizing the pipeline planning system at the national oil company, Eur. J. Oper. Res. (2019).
[91] D.C. Cafaro, J. Cerdá, Short-term operational planning of refined products pipelines, Optim. Eng. 18 (1) (2017) 241–268.
[92] A. Herrán, F. Defersha, M. Chen, J. Cruz, An integrated multi- period planning of the production and transportation of multiple petroleum products in a single pipeline system, Int. J. Ind. Eng. Comput. 2 (1) (2011) 19–44.
[93] W. Abdellaoui, M. Souier, M. Sahnoun, F.B. Abdelaziz, Multi-period optimal schedule of a multi-product pipeline: a case study in Algeria, Comput. Ind. Eng. 159 (2021) 107483.
[94] E.M. Souza Filho, L. Bahiense, L.S. de Lima, V.J.M. Ferreira Filho, Exact and heuristic approachs to the multiproduct pipeline scheduling problem, Anais do Simpósio Brasileiro de Pesquisa Operacional (SBPO) (2008).
[95] F. Khalili Goudarzi, H. Maleki, S. Niroomand, An optimization framework for scheduling multi-period multi-product oil pipeline systems under belief degree-based uncertain parameters, J. Appl. Math. Comput. 69 (1) (2023) 37–68.
[96] M. Bamoumen, S. Elfirdoussi, L. Ren, N. Tchernev, An efficient grasp-like algorithm for the multi-product straight pipeline scheduling problem, Comput. Oper. Res. 150 (2023) 106082.

[97] S.H. Moghimi, J. Habibi, H. Jahad, M.A. Fazli, Operational scheduling of oil products pipeline with intermediate event occurrences, arXiv preprint arXiv:2301.00451, 2023.
[98] N. Beheshti Asl, S. MirHassani, S. Relvas, F. Hooshmand, A novel two-phase decomposition-based algorithm to solve minlp pipeline scheduling problem, Oper. Res. 22 (5) (2022) 4829–4863.
[99] L. Magatão, L.V. Arruda, F. Neves-Jr, A combined clp-milp approach for scheduling commodities in a pipeline, J. Sched. 14 (1) (2011) 57–87.
[100] F.J. Moura Marcellino, O. Nizam, A.V. Moura, The planning of the oil derivatives transportation by pipelines as a distributed constraint optimization problem, in: Distributed Constraint Reasoning (DCR'07), 2007, p. 1.
[101] S. MirHassani, M. Ghorbanalizadeh, The multi-product pipeline scheduling system, Comput. Math. Appl. 56 (4) (2008) 891–897.
[102] Á. García-Sánchez, L.M. Arreche, M. Ortega-Mier, Combining simulation and tabu search for oil-derivatives pipeline scheduling, in: Metaheuristics for Scheduling in Industrial and Manufacturing Applications, Springer, 2008, pp. 301–325.
[103] H. Mostafaei, Y. Alipouri, M. Zadahmad, A mathematical model for scheduling of real-world tree-structured multi-product pipeline system, Math. Methods Oper. Res. 81 (1) (2015) 53–81.
[104] H. Chen, L. Zuo, C. Wu, Q. Li, An milp formulation for optimizing detailed schedules of a multiproduct pipeline network, Transp. Res., Part E, Logist. Transp. Rev. 123 (2019) 142–164.
[105] M. Taherkhani, An milp approach for scheduling of tree-like pipelines with dual purpose terminals, Oper. Res. (2018) 1–29.
[106] M. Taherkhani, M. Seifbarghy, R. Tavakkoli-Moghaddam, P. Fattahi, Mixed-integer linear programming model for tree-like pipeline scheduling problem with intermediate due dates on demands, Oper. Res. (2017) 1–27.
[107] S. MirHassani, An operational planning model for petroleum products logistics under uncertainty, Appl. Math. Comput. 196 (2) (2008) 744–751.
[108] Z. Li, Y. Liang, Q. Liao, N. Xu, J. Zheng, H. Zhang, Scheduling of a branched multiproduct pipeline system with robust inventory management, Comput. Ind. Eng. 162 (2021) 107760.
[109] A.V. Moura, C.C. de Souza, A.A. Cire, T.M. Lopes, Heuristics and constraint programming hybridizations for a real pipeline planning and scheduling problem, in: 11th IEEE International Conference on Computational Science and Engineering 2008 CSE'08, IEEE, 2008, pp. 455–462.
[110] S. Relvas, H.A. Matos, A.P.F. Barbosa-Povoa, J. Fialho, A.S. Pinheiro, Pipeline scheduling and inventory management of a multiproduct distribution oil system, Ind. Eng. Chem. Res. 45 (23) (2006) 7841–7855.
[111] D.C. Cafaro, J. Cerdá, Optimal scheduling of refined products pipelines with multiple sources, Ind. Eng. Chem. Res. 48 (14) (2009) 6675–6689.
[112] V.G. Cafaro, D.C. Cafaro, C.A. Méndez, J. Cerdá, Detailed scheduling of single-source pipelines with simultaneous deliveries to multiple offtake stations, Ind. Eng. Chem. Res. 51 (17) (2012) 6145–6165.
[113] D.C. Cafaro, J. Cerdá, Efficient tool for the scheduling of multiproduct pipelines and terminal operations, Ind. Eng. Chem. Res. 47 (24) (2008) 9941–9956.
[114] H. Mostafaei, A. Ghaffari Hadigheh, A general modeling framework for the long-term scheduling of multiproduct pipelines with delivery constraints, Ind. Eng. Chem. Res. 53 (17) (2014) 7029–7042.
[115] H. Chen, C. Wu, L. Zuo, F. Diao, L. Wang, D. Wang, B. Song, Optimization of detailed schedule for a multiproduct pipeline using a simulated annealing algorithm and heuristic rules, Ind. Eng. Chem. Res. 56 (17) (2017) 5092–5106.
[116] S. Relvas, H.A. Matos, A.P.F. Barbosa-Póvoa, J. Fialho, Reactive scheduling framework for a multiproduct pipeline with inventory management, Ind. Eng. Chem. Res. 46 (17) (2007) 5659–5672.
[117] D.C. Cafaro, J. Cerdá, A rigorous mathematical formulation for the scheduling of tree-structure pipeline networks, Ind. Eng. Chem. Res. 50 (9) (2011) 5064–5085.
[118] P.M. Castro, Optimal scheduling of pipeline systems with a resource- task network continuous-time formulation, Ind. Eng. Chem. Res. 49 (22) (2010) 11491–11505.
[119] V.G. Cafaro, D.C. Cafaro, C.A. Méndez, J. Cerdá, Detailed scheduling of operations in single-source refined products pipelines, Ind. Eng. Chem. Res. 50 (10) (2011) 6240–6259.
[120] H. Mostafaei, P.M. Castro, A. Ghaffari-Hadigheh, A novel monolithic milp framework for lot-sizing and scheduling of multiproduct treelike pipeline networks, Ind. Eng. Chem. Res. 54 (37) (2015) 9202–9221.
[121] S. MirHassani, S. Moradi, N. Taghinezhad, Algorithm for long-term scheduling of multiproduct pipelines, Ind. Eng. Chem. Res. 50 (24) (2011) 13899–13910.
[122] Q. Liao, P.M. Castro, Y. Liang, H. Zhang, Computationally efficient milp model for scheduling a branched multiproduct pipeline system, Ind. Eng. Chem. Res. (2019).
[123] M. Taherkhani, Discrete-time approach to operational scheduling of treelike pipelines with multiple input and output nodes, Ind. Eng. Chem. Res. 60 (49) (2021) 18018–18030.
[124] S.N. Boschetto, L. Magatão, W.M. Brondani, F. Neves-Jr, L.V. Arruda, A.P. Barbosa-Póvoa, S. Relvas, An operational scheduling model to product distribution through a pipeline network, Ind. Eng. Chem. Res. 49 (12) (2010) 5661–5682.
[125] P.M. Castro, Optimal scheduling of multiproduct pipelines in networks with reversible flow, Ind. Eng. Chem. Res. 56 (34) (2017) 9638–9656.
[126] S. Magatão, L. Magatão, F. Neves-Jr, L. Arruda, Novel milp decomposition approach for scheduling product distribution through a pipeline network, Ind. Eng. Chem. Res. 54 (18) (2015) 5077–5095.
[127] S.N.B. Magatão, L. Magatão, H. Luis Polli, F. Neves Jr, L.V.R. de Arruda, S. Relvas, A.P.F.D. Barbosa-Póvoa, Planning and sequencing product distribution in a real-world pipeline network: an milp decomposition approach, Ind. Eng. Chem. Res. 51 (12) (2012) 4591–4609.
[128] S.L. Stebel, S.N. Magatão, L.V. Arruda, F. Neves-Jr, A.P. Povoa, S. Relvas, Mixed integer linear programming formulation for aiding planning activities in a complex pipeline network, Ind. Eng. Chem. Res. 51 (35) (2012) 11417–11433.
[129] M. Taherkhani, R. Tavakkoli-Moghaddam, M. Seifbarghy, P. Fattahi, Detailed scheduling of tree-like pipeline networks with multiple refineries, Int. J. Eng. Trans. C: Aspects 30 (12) (2017) 1870–1878.
[130] L. Magatao, L. Arruda, F. Neves, Sequencing inputs to a multi-product pipeline, in: 2001 European Control Conference ECC 2001, IEEE, 2001, pp. 2152–2157.
[131] R. Tu, H. Zhang, S. Xu, G. Fu, Z. Li, Q. Liao, J. Du, Y. Liang, Pipeline sharing: boosting multi-product pipeline transport biofuels in the shift to low-carbon energy, J. Clean. Prod. 437 (2024) 140663.
[132] J. Krause, E.L. Sieczka, H.S. Lopes, Differential evolution variants and milp for the pipeline network schedule optimization problem, in: Computational Intelligence (LA-CCI), 2015 Latin America Congress on, IEEE, 2015, pp. 1–6.
[133] M.E. Baysal, E. Uçaktürk, Solution to a pipeline scheduling problem by using a mixed integer linear programming model, Selçuk Üniversitesi Mühendislik, Bilim ve Teknoloji Dergisi 4 (4) (2016) 284–296.
[134] R.K. Ahuja, T.L. Magnanti, J.B. Orlin, Network Flows: Theory, Algorithms, and Applications, Prentice Hall, 1993.
[135] R. Rejowski, J. Pinto, A rigorous minlp for the simultaneous scheduling and operation of multiproduct pipeline systems, in: European Symposium Computer Aided Process Engineering ESCAPE-15, vol. 20, 2005, pp. 1063–1068.
[136] E. Abbasi, V. Garousi, Multi-objective optimization of both pumping energy and maintenance costs in oil pipeline networks using genetic algorithms, in: IJCCI (ICEC), 2010, pp. 153–162.
[137] R. Techo, D. Holbrook, Computer scheduling worlds biggest product pipeline, Pipeline Gas J. 201 (5) (1974) 27–32.
[138] W. Abdellaoui, M. Souier, S. M'hammed, Modeling the scheduling of a multi products pipeline: a case study in Algeria, in: 2019 8th International Conference on Modeling Simulation and Applied Optimization (ICMSAO), IEEE, 2019, pp. 1–5.
[139] L. Magatão, L. Arruda, F. Neves Jr, A methodology for scheduling commodities in a multi-product pipeline, IFAC Proc. Vol. 35 (1) (2002) 205–216.
[140] H.-R. Zhang, Y.-T. Liang, Q. Liao, J. Ma, X.-H. Yan, An milp approach for detailed scheduling of oil depots along a multi-product pipeline, Pet. Sci. 14 (2) (2017) 434–458.

[141] D.S. Crane, R.L. Wainwright, D.A. Schoenefeld, Scheduling of multi-product fungible liquid pipelines using genetic algorithms, in: Proceedings of the 1999 ACM Symposium on Applied Computing, ACM, 1999, pp. 280–285.
[142] G.B. Dantzig, P. Wolfe, Decomposition principle for linear programs, Oper. Res. 8 (1) (1960) 101–111.
[143] M. Galati, T.K. Ralphs, Decomposition in integer linear programming, Ph.D. thesis, Lehigh University, 2009.