Computers and Chemical Engineering 150 (2021) 107319

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

Compact formulations for efficient early-phase field development optimization of multi-reservoir fields☆

Guowen Lei*, Thiago Lima Silva, Milan Stanko

Department of Geoscience and Petroleum, NTNU, 7031 Trondheim, Norway

ARTICLE INFO

Article history:
Received 27 November 2020
Revised 25 March 2021
Accepted 30 March 2021
Available online 3 April 2021

Keywords:
Field development optimization
Mixed-Integer linear programming
Piecewise-Linear functions

ABSTRACT

A compact formulation has been developed to efficiently optimize early-stage field development planning of multi-reservoir fields. The proposed formulation is a mixed-integer linear programming model which employs piecewise-linear functions to approximate the model non-linearities. The project economic value is maximized by optimizing the production allocation and the drilling schedule. The field production profiles are estimated with production potential curves calculated from simulated data of an integrated reservoir and surface facilities model. The novelties of this work are: a scalable model for the well combination selection, a logarithmic piecewise-linear model to approximate the well production potential curves, and the modeling and solution of realistic field development optimization problems. Through simulation analysis of a real field case study, the logarithmic and standard SOS2 formulations are compared in terms of computational performance and accuracy. The results show that the logarithmic formulation has significantly reduced the computational time and achieved improved accuracy over SOS2.

1. Introduction

The planning and development of oil and gas fields is a task of high complexity as it involves multiple disciplines and a large number of decisions. The field performance and project economics are highly dependent on the decisions made by the asset manager and the development planning team. In order to come up with a good development plan, it is important to consider a large range of possible scenarios involving the most relevant parameters during the planning phase. However, because there are often time constraints in the planning phase, the assessment of all possible scenarios is somewhat infeasible. For example, in Brazil, Exploration and Production (E&P) concession contracts stipulate that the decisions regarding the development of a field must be taken within the first 180 days after a commercial discovery is realized (Rodrigues et al., 2016). Therefore, in field planning, there is often the need to perform engineering calculations and sensitivity analysis to determine the most attractive design and quantify the effect of uncertain parameters within a limited time span.

Mathematical modeling was introduced to solve field development problems in the 1950s. To the best of our knowledge, Lee and Aronofsky (1958) were the first to publish a paper employing linear programming (LP) to solve the well drilling scheduling problem. After that, many works have been published in the literature on the use of mathematical programming methodologies to solve field development problems. A review of this literature can be found in (Durrer and Slater, 1977; Sullivan, 1988; Tavallali et al., 2016; Khor et al., 2017). The development of mathematical modeling applications in field development problems follows the advancing of computing speed and algorithmic techniques. Over all, the application of mathematical programming progressed from linear programming (1960 - 1980s) (Aronofsky and Williams, 1962; Attra et al., 1961; Lee and Aronofsky, 1958), to nonlinear programming (1960 - 1980s) (Rowan et al., 1967; McFarland et al., 1984), to mixed-integer linear programming (1970 - 2010s) (Rosenwald et al., 1974; Sullivan, 1988; Haugland et al., 1988; Nygreen et al., 1998; Iyer et al., 1998; Carvalho and Pinto, 2006), and the latest mixed-integer nonlinear programming (2000s - now) (Goel et al., 2006; Goel and Grossmann, 2004; Humphries and Haynes, 2015; Isebor et al., 2013; Lin and Floudas, 2003; van den Heever et al., 2001; Van Den Heever and Grossmann, 2000). Rosenwald et al. (1974) present a mathematical modeling proce-

☆ This research is a part of BRU21 – NTNU Research and Innovation Program on Digital and Automation Solutions for the Oil and Gas Industry (www.ntnu.edu/bru21).

* Corresponding author. .
E-mail addresses: guowen.lei@ntnu.no (G. Lei), thiago.l.silva@ntnu.no (T.L. Silva), milan.stanko@ntnu.no (M. Stanko).

https://doi.org/10.1016/j.compchemeng.2021.107319

dure using mixed-integer programming for determining the optimum location of wells assuming a set of pre-designed possible sites for new wells. In the paper surveys conducted in 1977 (Durrer and Slater, 1977), the problem of non-linearities of reservoir behavior and production network was described as difficult to handle. McFarland et al. (1984) used nonlinear programming techniques to solve field development planning and management optimization problems by selecting wells number, production rates, abandonment time and platform size. In their demonstrated cases, the well productivity is defined as a function of nonlinear pressure drop. In the work from Haugland et al. (1988), where they tested and presented the computational performance of field development optimization using mixed-integer programming, they concluded that the problem is hard to solve and the size of problems is limited by the computational capacity. Even today, many recently published papers report that mixed-integer nonlinear programming remains challenging due to its high computational requirements and unsatisfactory solution quality. The use of mathematical programming under decision-dependent uncertainty is another popular subject in recent years (Goel and Grossmann, 2004; Grossmann et al., 2016; Gupta and Grossmann, 2014a,b, 2017; Tarhan et al., 2009), but these problems are out of the scope of the present study.

In most studies, after the development of a model of the field value chain, some design parameters are optimized to improve some key performance indicators. The resulting problems are typically of large scale, non-linear, combinatorial, and combine black-box and analytical models. When formulating the optimization problem, there two main directions typically followed: minimization of the investment cost (Rodrigues et al., 2016; Grimmett et al., 1987; Devine and Lesso, 1972; Hansen et al., 1992; Garcia-Diaz et al., 1996), and maximization of the Net Present Value (NPV) (Frair and Devine, 1975; Huppler, 1974; Iyer et al., 1998; McFarland et al., 1984; Nygreen et al., 1998; Tavallali et al., 2013). In general, the methods targeting investment cost minimization deal with scheduling (e.g., drilling scheduling) and with how to place the platforms, wells, manifolds, pipelines and other relevant production facilities. Klose and Drexl (2005) reviewed and summarized the location and distribution problem in 2005, their results were extensively refereed and cited afterward. The methods that target the maximization of the NPV are typically focused on increasing the revenue and on cash flow analysis, mainly by improving the production planning and wells allocation.

In both investment cost minimization and NPV maximization problems, there are different ways to represent the decision variables and model their inter-relations. Different formulations to the same problem might be developed depending on the selection of the decision variables of the problem and of its main features. For instance, if it is desired to model the relation between the number of wells and the drilling sequence in the field performance, the influence of such decisions on the theoretical maximum production achievable must be included in the model, e.g., it is possible to produce higher rates if more wells are drilled. Iyer et al. (1998) proposed a list of the main decision variables involved in offshore field development problems:

(1) Number and location of production platforms, facilities and their capacities;
(2) Number and location of wells;
(3) Facilities (mainly platform and well) installation scheduling;
(4) Drilling rig location and scheduling;
(5) Production rate allocation for each time period.

In our work we are focusing on three variables from the list: the production rate allocation, the number of wells and the drilling scheduling. The production profiles over time are used to compute the revenue generation via cash flow analysis, and to compute the required capacity of processing facilities. The number of wells and the drilling scheduling define the maximum production rates that can be produced from the reservoir at each time step and this affects significantly the investment costs. According to statistics from the North Sea, the cost of well drilling accounts for about 40 - 50% of the total investment expenditure in offshore subsea field development projects (information compiled for the Norwegian Continental Shelf as of 2019) (Pavlov et al, 2020; NPD, 2020). As mentioned, the production allocation and the drilling schedule must be decided at an early stage of the field planning with limited and uncertain information. In subsequent stages of field development, these decisions are often frozen despite the availability of new information that could lead to improvements on the base design.

Production rate allocation is a process of allocating and forecasting recoverable reserves into a number of time periods or production horizons within the field's lifetime. This is typically performed with, e.g., three-dimensional reservoir models, decline curves (or type curves), material balance models and integrated coupled models of reservoir and surface network. Integrated coupled models of reservoir and surface network are often considered the most realistic because they capture adequately fluid energy losses from well bottom-hole to processing facilities. Rahmawati et al. (2012) evaluated optimal production strategies for an integrated field asset that coupled three reservoirs, a surface facility model and an economic model. Hepguler et al. (1997) present a study that couples a three-dimensional reservoir simulator with a general-purpose network simulator. Their study concludes that an integrated model gives a much more complete description of field behavior. However, it can be time-consuming and challenging to set up and run integrated models (Coats et al., 2003; Hepguler et al., 1997; Hoffmann et al., 2019). Some of the challenges are due to the complexity and non-linearity of the fundamental equations used to describe flow in reservoir and in surface network. Examples of optimization using a coupled model of reservoir-network are the works by Rahmawati et al. (2012), Hoffmann et al. (2019) and Silva et al. (2019). Hoffmann et al. (2019) proposed a solution that integrated the reservoir and network models built in commercial software. Silva et al. (2019) proposed the model integration based on an open-source fully implicit reservoir simulator such that the gradients are made available to the optimization algorithm through automatic differentiation. The well and network models are described with mechanistic equations based on physical principles.

Despite being more accurate, when it comes to optimization, the computational time required for achieving a solution using non-linear coupled reservoir-network models is usually prohibitive. Haugland et al. (1988) compared the computational performance - CPU time - in terms of numbers of integer and continuous variables, the computation is extremely time-consuming even with only 23 integer variables and 60 continuous variables, demonstrating that problems including integer decision variables are hard to solve. Furthermore, models are often created using commercial software that are black-box, which significantly limits the availability and effectiveness of optimization tools. Therefore, many research studies have been focusing on the development of methods that utilize reduced computational resources. For instance, Sullivan (1988) discussed and illustrated a method to convert implicit production behavior to explicit models to solve much larger problems effectively. Goel et al. (2006) present a dual Lagrangian-based branch-and-bound algorithm to achieve a significant reduction in the model size. Gupta and Grossmann (2012) compared different solvers used in the optimization, and pointed out that reformulating MINLP into an MILP can solve the problem in an efficient way.

An alternative to full-fledged models is the use of low-order and reduced-complexity models, also known as proxy models. Because of its computational advantages, proxy models are often used to perform production scheduling with varying degrees of complexity and accuracy. A simple approach, for example, is assigning a production profile to each well. With such proxies, it is possible to model the drilling schedule, where wells are drilled at different points in time, but the well inter-dependency with varying target rates is not captured in the field performance. An example of this approach is the work of Wang et al. (2019), where they apply linear superposition of base production curves for each producer. Such methods are often used in facility placement and routing optimization problems since the main focus is on minimizing the investment cost, see Rodrigues et al. (2016).

Another approach to model production scheduling is a hybrid method where proxy models are used only for specific parts of the production system. For example, in the work by Iyer Iyer et al. (1998) and Van Den Heever and Grossmann (2000), reservoir pressure, producing gas-oil ratio and water cut are represented as a non-linear function of the cumulative production, which is extracted from the output of a reservoir simulation. Well deliverability, pressure drop in wellbore and flowlines were solved using mechanistic equations based on physical principles. Iyer et al. (1998) proposed a sequential decomposition strategy using aggregation and disaggregation technique for the planning and scheduling of investment and operation in offshore oil field facilities. Carvalho and Pinto (2006) used the algorithm of Iyer et al. on an offshore oilfield infrastructure planning problem but modifying the branching priorities and solver parameters to decrease the computation cost. Lin and Floudas (2003) used a similar method to characterize the non-linear reservoir performance of gas fields and included it into the well platform planning problem.

A field-level proxy method that allows to represent the well inter-dependency and variations in the wells and field target rates are the production potential curves. The production potential curves can be seen as an upper bound to the oil or gas rates, which are feasible to be produced by the production system at a specific depletion state. At any given time, one can decide to produce at the potential or at any value below it. Goel et al. (2006) and Goel and Grossmann (2004) assume a linear relationship between the field's deliverability and the recovered amount of hydrocarbons. By changing the end points of the curve they represented the uncertainties in reservoir size and field productivity.

González et al. (2019) and Angga (2019) performed mixed integer-piecewise linear optimization of drilling and production scheduling in early-phase field planning using a collection of production potential curves. The curves were generated with a coupled reservoir-network model and depended on the number of wells, reservoir size, network layout, artificial lift mechanism and reservoir recovery method. One of the drawbacks of their approach is that the process to generate production potential curves might be time-consuming when there are many variables to consider.

Stanko (2021) further developed this method and introduced the use of dimensionless production potential curves by dividing cumulative production values by reservoir size and production potential by their upper bound. Stanko showed that, in many cases, the dimensionless production potential curve is not affected significantly by network layout, heterogeneity in the well and gathering network, well count, reservoir size, separator pressure and artificial lift method. Therefore, in such cases, production potential curves can be generated by scaling dimensionless production potential curves with the maximum production potential and reservoir size of the case. Alkindira (2020) used this scaling technique in an early-phase field development optimization problem while considering uncertainties in-place volumes and scheduling of wells with distinct performance.

Special Ordered Set 2 (SOS2) is used to piecewise-linearize production potential curves yielding Mixed-Integer Linear Programming formulations (Angga, 2019; Alkindira, 2020; González et al., 2019). This approach of using Special Ordered Sets on piecewise linearization has been used extensively in the past by e.g., Sullivan (1988); Hoffmann et al. (2019); Gupta and Grossmann (2012); Gunnerud et al. (2012); Gunnerud and Foss (2010); Epelle and Gerogiorgis (2020); Rosa et al. (2018), often to approximate the well and pipe flow behavior for a hydrocarbon production system. For instance, Sullivan (1988) used Special Ordered Sets (SOS) to identify reservoir production alternatives when formulating the optimization problem as an MIP model. Epelle and Gerogiorgis (2020) used SOS2 to piecewise-linearize the pressure-rate responses in a production system encompassing the wells, routing in the gathering network and pipelines. In this paper, they did an extensive computational analysis to compare MILP and MINLP formulations through 3 case studies. All these studies show improvements in computational efficiency after using SOS2 to convert the Mixed-Integer Non-Linear Programming (MINLP) problems into Mixed-Integer Linear Programming (MILP) problems. However, other studies indicated that in more complex cases the performance of SOS2 might not be satisfactory. Silva and Camponogara (2014) compared different algorithms used to piecewise-linearize a gas lift optimization problem. In their study, it is shown that even though the formulation using SOS2 variables can find the global optimal, it struggles when applied to complex problems involving fine multidimensional nonlinear approximations. This is consistent with the observations provided by Brito et al. (2020) in a hydro unit commitment problem, where they concluded that SOS2 is not as efficient as the Logarithmic Convex Combination (Log) model (Vielma et al., 2010).

1.1. Paper objective

In this paper, we formulate the early-phase offshore oil field development problem with a mathematical programming model focusing on production allocation and well drilling scheduling. We first list all relevant well combinations and compute their maximum production potential using an integrated reservoir-facility model. In our mathematical model, we propose a novel formulation for the selection of the well combination that is scalable to fields with a large number of wells because only one binary variable per well is required. We propose the use of Log, a compact piecewise-linear model, to approximate the non-linear functions representing the production potential, the field water and gas production. The efficiency and accuracy of the resulting MILP formulation with Log are assessed in a case study of a real hydrocarbon field, where the results obtained with the proposed formulation are compared against standard SOS2-based models. The distinct features of our work are as follow:

(1) We propose a formulation for the well combination selection and activation that is easy to expand to fields with a large number of wells as it only requires one binary variable per well.
(2) Although the Log model has been used in the past in production optimization problems to approximate pressure drop and well production curves Silva and Camponogara (2014), to the best of our knowledge, despite its promising computational efficiency, our work is the first to employ a Logarithmic-based piecewise-linear model to represent the nonlinear functions of field development optimization problems. In this work, detailed piecewise-linear formulations of the standard SOS2 and the Log model are presented for a field development optimization problem of considerable complexity.
(3) After a comprehensive comparison between the standard SOS2 model and the proposed Logarithmic model, We

demonstrate using a real-world case study that the optimization with Log is considerably more efficient, mainly because of its compactness and the strength of the resulting relaxations, as it was also described by Vielma and Nemhauser (2011). Furthermore, the proposed approximations with Log have improved accuracy compared to standard SOS2 techniques because of the higher resolution approximations yielded through the use of simplices (instead of hypercubes) in the domain partitioning.

(4) We demonstrate the effectiveness of the new model for field development optimization problems, which combines a scalable formulation for well combination selection and a Logarithmic model for the piecewise-linear approximations, by applying it on a real-world case study of an early-phase offshore field development planning problem. The case investigated in the paper is more realistic than others found in the literature, which can also be regarded as a contribution in terms of modeling and application in the field development research area.

1.2. Paper structure

This paper is organized as follows. Section 2 and Section 3 present the problem description and a mathematical model, respectively. Then, in Section 4, an approximate mixed-integer linear programming model is proposed using piecewise linear (PWL) models. In Section 5, we demonstrate the advantages of our approach with a set of illustrative examples and a real-world case study. The conclusions are presented in the last section of the paper.

2. Field development optimization

In offshore field development projects, it is often common to commingle the production of multiple neighboring reservoirs into the same platform and facilities. These multi-reservoir fields are become more and more common not only because of technological advances but also for economic reasons, i.e., in some cases is not profitable to develop a small size reservoir independently. Multi-reservoir fields can also be developed when the new discovery is made in nearby regions of mature fields, which often have an extra capacity of processing and transportation. In either of the cases, the oil company chooses the most financially beneficial development concept that can including all discovered reservoirs in their asset region.

2.1. Problem statement

The case study for this paper is a field with two reservoir units, subsea wells and gathering network producing to a production platform as illustrated in Fig. 1. The production from the wells in each reservoir unit goes through their wellhead to a subsea manifold, where it is commingled into a pipeline network. Reservoir 1 has a total of 6 wells and Reservoir 2 has a total of 3 wells.

The variables to determine using mathematical maximization of project value are production rate allocation per year, the total number of wells required in each reservoir and drilling scheduling. More wells and higher production increase the revenue stream due to hydrocarbon sales, but they also increase the operational and capital expenditures, e.g., topside facilities need to be bigger and more wells cost more.

For a given reservoir size, the maximum producible reserves (or the fraction of the hydrocarbon initial in place) are fixed depending both on the recovering strategy and production mechanism. In this paper, we assume the reservoir production is driven by natural depletion, which means that no secondary recovery mechanism such as water injection or gas injection is employed. We also assume that the location of the production platform, the well-heads, the manifolds are known and fixed a priori and have been determined by layout optimization methodologies and seabed geological survey studies. Additionally, we assume that the initial oil in place of both reservoirs is deterministic and known. Furthermore, the following extra assumptions and considerations are made:

(1) There is no underground flow communication between the reservoirs.
(2) The production from reservoir 1 is hydraulically decoupled from the production from reservoir 2.
(3) The field's main product is oil, but it also produces some associated gas and water. The producing gas-oil ratio and water cut are a function of cumulative oil production.
(4) The production performance is unique for each well and for the overall field for different well combinations.
(5) Production potential curves are used to define the upper bound of production profiles.

The usage of production potential curves to constrain the optimal production rates is widely adopted, as seen in previous works from Gupta and Grossmann (2012, 2017); Lin and Floudas (2003); Goel et al. (2006); Goel and Grossmann (2004); Tarhan et al. (2009); González et al. (2019); Stanko (2021). Basically, it is a numerical representation based on the material balance of the production performance of a production system. The production system can comprise of reservoir and wells or an integrated system including reservoir, wells and gathering network to the processing facilities. It is derived from the rate-pressure-volume relation to a rate vs. cumulative production or recovery factor. Curve's shape in linear or nonlinear, convex or non-convex indicates the degree of complexity and understanding of the system.

In this paper, production potential curves (field production potential versus oil cumulative production) were extracted for each reservoir and subsea system from simulations of coupled reservoir-network models. The simulations consider all wells in each reservoir are active and produce as much as possible. To determine the production potential when only a particular combination of wells is active, we use a variation of the method described by Stanko (2021): 1) perform steady-state well and network model simulations of the particular well combination w_j at initial time and record rate $q^r_{\text{pot,max}}(w_j)$ of reservoir r; 2) scale the production potential values of the curve by the factor $fn = q^r_{\text{pot,max}}(w_j)/q^{r,\text{all}}_{\text{pot,max}}$, where $q^{r,\text{all}}_{\text{pot,max}}$ is the production potential of the curve considering all wells from reservoir r are active at initial conditions. This procedure assumes that the curve of current dimensionless production potential of a given reservoir and subsea system is not affected significantly by well combination. Please note that our model does not track how much is produced by each well, it only considers what a set of wells produces. Therefore, in our formulation, it is usually not possible to apply rate constraints on a well level, only on a field level (unless there is only one well producing). Nevertheless, different field production potentials are considered depending on which wells are drilled and produced from (well combination) over the field life time.

Fig. 2 shows the field production potential of reservoir 1 with 3 producers and for 3 distinct well combinations. As it can be seen, the field production potential is scaled up or down depending on the well combination. The combination of wells $w3, w4, w5$ gives higher production than $w1, w4, w5$ and $w1, w2, w3$. However, towards the right of the curve, all wells combinations converge to similar values of oil cumulative production. Using wells $w3, w4, w5$ allows to produce higher rates at early times, but the processing capacity of the production platform must also be increased, which costs more and may reduce the value of the project.

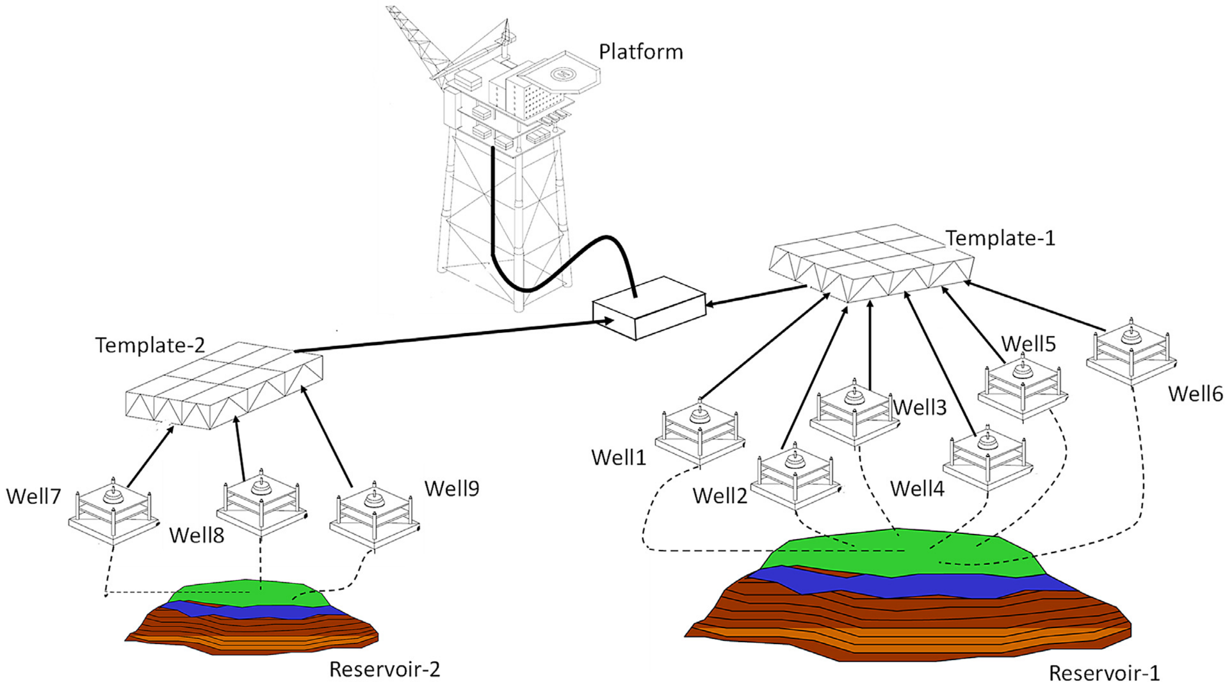


Fig. 1. Field layout.

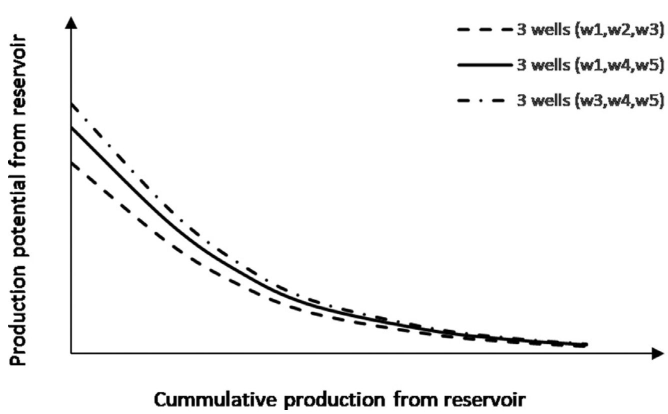


Fig. 2. Curves of production potential versus cumulative production of a reservoir using 3 distinct well combinations.

The produced gas and water rates of each reservoir are forecasted from curves of cumulative gas production and cumulative water production versus cumulative oil production. This approach assumes that GOR and WC are functions of cumulative oil production only and are not affected by well combination. However, the arithmetic operations between oil rate, GOR and WC to obtain gas and water rates are non-linear, and must then be linearized to be compatible with a MILP formulation. Angga (2019) compared two methods to compute water and gas rates: 1) a bi-linearization of the arithmetic operations and 2) computing oil and gas rates from curves of cumulative gas production and cumulative water production versus cumulative oil production. He showed that the latter approach is significantly more computationally efficient. A similar observation is provided by Gupta and Grossmann (2012). The produced gas and water rates also impact the capacity, design and ultimately the cost of topside facilities.

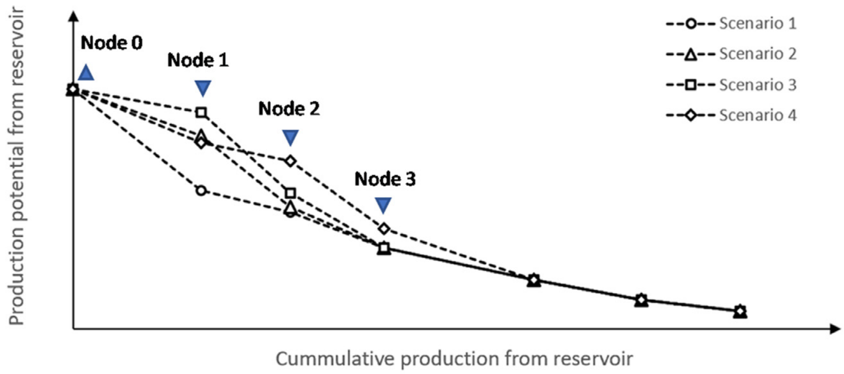
In addition to the well combination, the drilling sequence also impacts the field production potential. Fig. 3 illustrates the production potential curve of Reservoir 1 where the well combination is varied at specific cumulative oil production values, i.e., specific wells are drilled and start production at each point. At initial time, wells $w1, w2$ and $w3$ are active. The production potential curve will change when one chooses different wells to be drilled in each decision node. If, for example, one wishes to produce the field at an oil high plateau rate, scenario 3 gives the longest plateau duration from all scenarios. If, alternatively, one wishes to produce the field at a long oil plateau rate, scenario 4 gives the longest plateau duration from all scenarios.

The drilling schedule and well combination affect not only the production profile but also the investment profile (drilling expenditures and facilities expenditures). Thus, the production schedule, drilling schedule and well combination must be determined such that the overall economic value of the project is maximized. In this work we use the Net Present Value (NPV) as an economic indicator, which includes the discounted revenue obtained with the products' sales and the overall expenditures required to develop the field.

3. Mathematical formulation

In this section we present a mathematical programming model for multi-reservoir field development optimization. The objective is to maximize the net present value of the project and the main decision variables are the drilling and production schedule. We introduce a novel approach to determine the well combination and the total number of active wells over the field producing time. The proposed formulation is as generic as possible, such that it will be possible to expand it to deal with more reservoir units, adding more wells, studying complex drilling scenarios, and including different cost functions, among others. Several equality and inequality constraints are formulated to represent physical limitations of the system. For instance, the field production is constrained by the process capacity limitations, the number of drilling wells is restricted by the maximum drilling capacity (e.g., the window to drill wells in Artic areas is usually less than 5 months per year).

The description of the model is divided into notation, objective function, and constraints. The following section introduces the



Scenario	Decision Node 0	Decision Node 1	Decision Node 2	Decision Node 3
1	(w1,w2,w3)	(w1,w2,w3)	(w1,w2,w3,w4)	(w1,w2,w3,w4,w5)
2	(w1,w2,w3)	(w1,w2,w3,w5)	(w1,w2,w3,w5)	(w1,w2,w3,w4,w5)
3	(w1,w2,w3)	(w1,w2,w3,w4,w5)	(w1,w2,w3,w4,w5)	(w1,w2,w3,w4,w5)
4	(w1,w2,w3)	(w1,w2,w3,w4)	(w1,w2,w3,w4,w5,w6)	(w1,w2,w3,w4,w5,w6)

Fig. 3. Curves of production potential versus cumulative production when different well combinations are enforced at specific values of cumulative production .

Table 1
Sets and indices.

$\mathcal{T}$	Set of all time steps
$\mathcal{R}$	Set of reservoirs
$\mathcal{W}^r$	Set of wells in reservoir r
$\mathcal{KG}$	Set of breakpoints in the cumulative gas production(Gp)
$\mathcal{KW}$	Set of breakpoints in the cumulative water production(Wp)
$\mathcal{KF}$	Set of oil rate breakpoints q_o in the domain of the potential Np_{pot1}
$\mathcal{KQ}$	Set of oil rate breakpoints q_o in the domain of the function fn_{pot2}
(i)	Well indices $i \in \{1,\dots,W^r\}$
(j)	Well permutations $j \in \{1,\dots,2^{W^r}\}$

Table 2
Continuous variables.

$q_o(t)$	Oil production in period t
$q_g(t)$	Gas production in period t
$q_w(t)$	Water production in period t

Table 3
Integer Variables.

$N_w^r \in \mathbb{Z}$	Number of wells of reservoir r
$N_w^f \in \mathbb{Z}$	Number of wells of field f
$x_i^r \in [0,1]$	Status of well i in reservoir r

model notation. Section 3.2 presents the objective function, and Section 3.3 presents the model constraints.

3.1. Notation

All the sets and indices used in the mathematical formulation are presented in Table 1. The continuous and integer variables can be found in Tables 2 and 3, respectively. The parameters used in the mathematical formulation are shown in Table 4. Finally, the superscripts utilized in the formulation appear in Table 5.

3.2. Objective function

The objective function to be maximized is the Net Present Value (*NPV*) formulated in Eq. (1), which is the sum of yearly cash flows discounted to time "zero". The yearly cash flow includes the revenue obtained from oil and gas sales subtracted by the costs of investment in facilities, drilling and operation of the field:

$$\max\ NPV = \sum_{t}^{T} \frac{Revenue^f(t) - Cost^f(t)}{(1+D)^t} \tag{1}$$

$$Revenue^f(t) = P_o(t) \times q_o^f(t) + P_g(t) \times q_g^f(t) \tag{2}$$

$$Cost^f(t) = CAPEX^f(t) + OPEX^f(t) \tag{3}$$

The yearly cash flow is discounted to its present value using the discount factor D, which is a decimal number. The commodity price P_o and P_g are used as inputs in the revenue calculation Eq. (2). We assume the commodity price is constant during the lifetime of the field, but the formulation can be extended to consider a varying commodity price. The cost is split into capital expenditure-*CAPEX*(the cost associated with drilling, facilities construction and installation, etc.) and operation expenditure-*OPEX*(the cost associated with production operations) in Eq. (3).

3.3. Constraints

In this section we present the model constraints. The constraints are split into production rate, cumulative production, well, CAPEX, OPEX and well scheduling constraints.

3.3.1. Production rate constraints

Assuming that reservoirs $r \in \mathcal{R}$ in the field f are independent, the total oil, gas and water production in the field are calculated as the sum of the production coming from all reservoirs, as formulated in Eqs. (4), (5) and (6). These total production rates are constrained by the respective capacities of the processing facilities at the production platform, as stated in Eqs. (7), (8) and (9), where $q_o^f(t)$, $q_g^f(t)$, and $q_w^f(t)$ are respectively the total oil, gas, and water rates produced by the field, which are bounded by the corresponding topside capacities q_o^{max}, q_g^{max}, and q_w^{max}. The field's production rate must be selected such that the revenue due to hydrocarbon sales out-weights the costs of facilities, see Jahn et al. (2008).

The third constraint is the production potential curve. The oil production rate of each reservoir r is bounded by its production potential at a given point in time in Eq. (10), where $q_0^r(t)$ is a variable denoting the production rate from reservoir r at time/year t,

Table 4
Parameters.

D	Discount factor
q_o^{max}	Maximum oil rate in the production platform
q_g^{max}	Maximum gas rate in the production platform
q_w^{max}	Maximum water rate in the production platform
$N_w^{f,Start}$	Pre-drilled well in field f
N_w^{Dmax}	Maximum drilling capacity per year
N_D	Total number of years in which the initial CAPEX is distributed
L_{pipe}^f	Length of pipeline
N_{joint}^f	Number of subsea joints-template,manifold,pump etc.
P_o	Oil price
P_g	Gas price
α_1	$CAPEX_{Drilling}$ linear coefficient of single well drilling expenditure
α_2, α_3	$CAPEX_{Subsea}$ linear coefficient of pipeline length expenditure & subsea joint expenditure
$\alpha_4, \alpha_5, \alpha_6$	$CAPEX_{Topside}$ linear coefficient of maximum oil, gas and water processing capacity
$\alpha_7, \alpha_8, \alpha_9$	$OPEX_{rate}$ linear coefficient of the oil, gas and water rate
$\alpha_{10}, \alpha_{11}, \alpha_{12}$	$OPEX_{Nonrate}$ linear coefficient of the well number, pipeline length and joints number
$\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$	Constant term in the linear function of $CAPEX_{Drilling}$, $CAPEX_{Subsea}$, $CAPEX_{Topside}$, $OPEX_{rate}$ and $OPEX_{Nonrate}$

Table 5
Superscripts.

(f)	Field
(r)	Variables associated with reservoir r $\in \mathcal{R}$
(o)	Oil
(g)	Gas
(w)	Water
(t)	Time periods

and $q_{o,pot}^r(t)$ is the production potential of reservoir r at time/year t. Eq. (10) ensures that the production does not exceed the maximum feasible oil production rate of the reservoir.

$$q_o^f(t) = \sum_{r=1}^{R} q_o^r(t) \tag{4}$$

$$q_g^f(t) = \sum_{r=1}^{R} q_g^r(t) \tag{5}$$

$$q_w^f(t) = \sum_{r=1}^{R} q_w^r(t) \tag{6}$$

$$q_o^f(t) \le q_o^{\max} \tag{7}$$

$$q_g^f(t) \le q_g^{\max} \tag{8}$$

$$q_w^f(t) \le q_w^{\max} \tag{9}$$

$$q_o^r(t) \le q_{o,pot}^r(t) \tag{10}$$

3.3.2. Cumulative production constraints

The cumulative production of oil, gas, and water for each reservoir are calculated based on the cumulative production of the previous time step or at the start of the production phase, and the production in the previous time step (backward approximation). In order to simplify the calculations, the time step is assumed to be 1 year, and the unit used for the cumulative production N_p, G_p and W_p is $10^3\,Sm^3$, whereas the unit for the oil (q_o), water (q_w) and gas rate (q_g) is $10^3\,Sm^3/Year$. We chose a time step of 1 year because of 4 reasons: 1. In early phases of field development, the industry typically performs discounted cash flow calculations on a yearly basis; 2. Most past and recent previous works in the literature also use a time step of a year (e.g. Epelle and Gerogiorgis, 2019; Wang et al., 2019; (Gupta and Grossmann, 2012); 3. Our problem has some constraints that only make sense in a time frame of a year, for example the maximum number of wells that can be drilled in a year; 4. Decreasing the time step length will increase dramatically the running time of the model and will make it challenging to run uncertainty analyses with it. Both the gas and the water rates are back calculated from the cumulative gas and water production, which are a function of the actual cumulative oil production, as formulated in Eqs. (13) and (15).

The cumulative oil production from reservoir r is defined as:

$$N_p^r(t) = N_p^r(t-1) + q_o^r(t-1), t \ge 1 \tag{11}$$

$$N_p^r(0) = 0, \tag{12}$$

as the cumulative gas production from reservoir r is:

$$G_p^r(t) = G_p^r(t-1) + q_g^r(t-1), t \ge 1 \tag{13}$$

$$G_p^r(0) = 0 \tag{14}$$

and finally the cumulative water production from reservoir r is defined as:

$$W_p^r(t) = W_p^r(t-1) + q_w^r(t-1), t \ge 1 \tag{15}$$

$$W_p^r(0) = 0 \tag{16}$$

3.3.3. Well constraints

The wells activation and the corresponding number of drilled wells in each time step t are formulated with the following equality constraint of Eq. (17). The total number of wells in reservoir r is $N_w^r(t)$ and the total number of well of the field is denoted by $N_w^f(t)$ in Eq. (18):

$$N_w^r(t) = \sum_{i}^{W^r} x_i^r(t) \tag{17}$$

$$N_w^f(t) = \sum_{r=1}^{R} N_w^r(t) \tag{18}$$

where $x_i^r(t)$ is a binary variable representing the well status, which takes on value 1 in case the well is opened at time step t, and 0 otherwise.

The number of wells is constrained by physical limitations both at the reservoir and field levels, e.g., the maximum number of wells allowed to be connected to the template and production platform (Eqs. (19) and (20)). In some types of offshore production structures, such as steel jackets and SPAR floaters, there is usually a limited number of well slots available on the deck.

$$N_w^f(0) \leq N_w^f(t) \leq N_w^{f,max}, t \geq 1 \tag{19}$$

$$N_w^r(0) \leq N_w^r(t) \leq N_w^{r,max}, t \geq 1 \tag{20}$$

The number of pre-drilled wells is defined as $N_w^{f,Start}$, and it is equal to the number of wells at the beginning of the first year of the production, when $t = 0$, as formulated in Eq. (21). Once the production starts, there is a limit on the number of yearly wells that can be drilled in the field N_w^{Dmax} (see Eq. (22)), e.g., no more than 3 wells can be drilled per year. Eq. (23) ensures that, after a well is drilled, it can not be "shut-in" (un-drilled or abandoned).

$$N_w^f(0) = N_w^{f,Start} = \sum_{r=1}^{R} N_w^r(0), t = 0 \tag{21}$$

$$0 \leq N_w^f(t) - N_w^f(t-1) \leq N_w^{Dmax}, t \geq 1 \tag{22}$$

$$x_i^r(t+1) \geq x_i^r(t) \tag{23}$$

In some cases, pre-drilling some wells before producing the first oil improves the project cash flow in an offshore oil/gas field, see Jahn et al. (2008).

3.3.4. CAPEX constraints

Linear equations were employed in the cost models for drilling, facilities, and operational expenditures. Despite being linear, the model is flexible and allows extensions to more complex cost models. The cost model depends on the yearly and maximum oil, gas and water flow rates and the number of wells.

The field CAPEX costs are be defined as:

$$CAPEX^f(t) = CAPEX_{Drilling}(t) + CAPEX_{Subsea}(t) + CAPEX_{Topside}(t) \tag{24}$$

where $CAPEX_{Drilling}(t)$ is the drilling cost, which is a function of the number of wells N_w^f drilled at a given time t. $CAPEX_{Subsea}(t)$ includes the costs of pipelines, manifolds and any other subsea layout structures. The cost of topside facilities $CAPEX_{Topside}(t)$ is a function of the maximum installed capacity for processing oil, gas and water rates. All expenditure of the facilities' fabrication and installation can be allocated to $CAPEX_{Subsea}(t)$ or $CAPEX_{Topside}(t)$.

The drilling CAPEX is defined as a linear relation of the number of drilled wells N_w^f at a given time t multiplied by the expenditure of drilling a single well α_1:

$$CAPEX_{Drilling}(t) = \alpha_1 \times (N_w^f(t) - N_w^f(t-1)) + \beta_1 \tag{25}$$

where β_1 are is a constant. Wellhead costs can be included into the parameter α_1 as they are proportional to the number of wells.

The subsea facilities CAPEX are defined as:

$$CAPEX_{Subsea}(t) = \frac{\alpha_2 \times L_{pipe}^f + \alpha_3 \times N_{joint}^f + \beta_2}{N_D}, \forall t \in \{1, \ldots, N_D\} \tag{26}$$

Notice that the joints can be regarded as manifolds that connect the wells and the pipes, but also flowline joints or subsea pumps. The length of the pipelines are defined as L_{pipe}^f, and N_D is the total number of years in which the initial CAPEX is uniformly distributed. The value of N_D partly depends on the tax regulation of the country and may vary from project to project. Values of 3 - 4 are common in the North Sea. L_{pipe}^f and N_{joint}^f can be defined as variables or parameters, depending on the particular case.

The topside CAPEX costs are modeled as a function of the designed maximum rates of oil (q_o^{max}), gas (q_g^{max}) and water (q_w^{max}) at the processing unit:

$$CAPEX_{Topside}(t) = \frac{\alpha_4 \times q_o^{max} + \alpha_5 \times q_g^{max} + \alpha_6 \times q_w^{max} + \beta_3}{N_D}, \quad \forall t \in \{1, \ldots, N_D\} \tag{27}$$

3.3.5. OPEX constraints

The operation costs (OPEX) can be divided into rate-dependent costs and non-rate costs:

$$OPEX^f(t) = OPEX_{rate}^f(t) + OPEX_{Nonrate}^f(t) \tag{28}$$

with the rate-dependent OPEX being defined as:

$$OPEX_{rate}^f(t) = \alpha_7 \times q_o^f(t) + \alpha_8 \times q_g^f(t) + \alpha_9 \times q_w^f(t) + \beta_4 \tag{29}$$

and the non-rate OPEX as:

$$OPEX_{Nonrate}^f(t) = \alpha_{10} \times N_w^f(t) + \alpha_{11} \times L_{pipe}^f + \alpha_{12} \times N_{joint}^f + \beta_5 \tag{30}$$

The rate-dependent OPEX is a function of the oil, gas, and water rates, whereas the non-rate costs are not. For rate-related costs, usually higher production rates lead to increased operational costs. Non-rate costs are typically costs involved in operations of maintenance, inspections and offshore personnel, transport, insurance. They are often dependent on the number of wells, the length of the pipelines and the subsea layout.

3.3.6. Well scheduling and status

In the formulation, the production potential depends on the active wells in the field, i.e., the well combination. We decided not to track all combinations (and field potentials) by assigning one binary variable per combination because the number of possible combinations grows exponentially with the number of wells, making the method non-scalable, i.e., the computational cost will be prohibitive for large systems. Instead, we utilize mapping and a set of disjunctions that yield a formulation which requires only one binary variable per well but still accounts for the different potentials of each well permutation. The compactness of the formulation contributes to its computational efficiency and also allows dealing with early-phase field development planning for longer producing time and a larger number of wells.

In order to account for the effect of well combinations on the production potential, we define the production potential for each reservoir r at time t as follows:

$$q_{o,pot}^r(t) = fn^r(t) \times f_q(N_p^r(t)) \tag{31}$$

where $fn^r(t)$ is a factor that varies continuously in the interval [0,1] that indicates the actual production potential of the field r for a selected subset of producing wells among all the possible well permutations $j \in \{1, \ldots, 2^{|\mathcal{W}^r|}\}$. The actual production of reservoir r also depends on the maximum production potential $f_q^r(N_p^r(t))$, which is a function of the cumulative oil production N_p of reservoir r at time t.

The well status is represented by the variable $x_i^r \in \{0, 1\}$, where 0 means that the well is shut-in, and 1 means that it is producing. The well permutation is therefore denoted by a tuple $w_j^r = < x_1^r, x_2^r, \ldots, x_n^r >$ describing the status of the wells $i \in \{1, \ldots, n\}$ with n being the total number of wells (active and/or inactive) in reservoir $r \in \mathcal{R}$.

One straightforward way to model the selection of a well combination from all possible permutations w_j^r of reservoir r is to assign one binary variable per permutation j. However, this would require a total of $2^{|\mathcal{W}^r|}$ binary variables, which can be intractable for fields with a large number of wells. As the number of binary variables in a mixed-integer formulation affects considerably the computational time to obtain optimal solutions, we propose a novel modeling approach that requires only $|\mathcal{W}^r|$ binary variables to model the well combination selection.

We start by defining a function $g^r : \{1, \ldots, 2^{|\mathcal{W}^r|}\} \rightarrow \mathcal{W}^r$ that maps an index j from all permutations of well combinations w^r_j to the set of wells $\mathcal{W}^r$ such that:

$$g^r(j) = \{i \in \mathcal{W}^r : w^r_j(i) = 1\},\ \forall j \in \{1, \ldots, 2^{|\mathcal{W}^r|}\} \quad (32)$$

where $w^r_j(i)$ denotes the i-th element of the tuple w^r_j. This function indicates which wells $i \in \mathcal{W}^r$ are active in reservoir r for the j-th well combination. A table with a map generated by function $g^r(\cdot)$ is calculated off-line and used in the constraints regarding the well combination selection as follows:

For all $r \in \mathcal{R}, j \in \{1, \ldots, 2^{|\mathcal{W}^r|}\}$:

$$fn^r(t) \leq fn^r_j + \sum_{i \in g^r(j)} (1 - x^r_i(t)) + \sum_{i \in \mathcal{W}^r \setminus g^r(j)} x^r_i(t),\ \forall i \in \{1, \ldots, n\} \quad (33)$$

$$fn^r(t) \geq fn^r_j - \sum_{i \in g^r(j)} (1 - x^r_i(t)) - \sum_{i \in \mathcal{W}^r \setminus g^r(j)} x^r_i(t),\ \forall i \in \{1, \ldots, n\} \quad (34)$$

Eqs. (33) and (34) create a set of disjunctions such that, depending on the selection of the active wells through the binary variables x^r_i, the potential factor of the reservoir $fn^r(t)$ will be set to the potential factor fn^r_j corresponding to the correct well combination w^r_j from all permutations.

4. Piecewise-Linear approximations

The field development optimization problem formulated with Eqs. (1) – (34) is a Mixed-Integer Non-Linear Programming (MINLP) problem. It is mixed-integer because it contains both continuous variables regarding the well and field rates, and integer variables related to the status of the wells and the number of drilled wells. The nonlinearities of the problem appear in the production potential curves, including the actual potential based on the wells permutation, but also in the cumulative production rates for all the phases. The presence of discrete variables combined with the non-linear curves makes the optimization problem hard to solve. Our approach is to transform the MINLP problem into a Mixed-Integer Linear Programming (MILP) one by utilizing Piecewise-Linear (PWL) functions to approximate the non-linearities, which is similar to the approach used in Silva and Camponogara (2014) for production optimization problems.

In the optimization problem formulated in this work, there are a total of 3 non-linear functions, which are actually not available in an explicit form, and will be sampled from simulations and interpolated with PWL functions. Among such functions are the cumulative gas G^r_p and water production W^r_p, which are a one-dimensional function of the cumulative oil production N^r_p. The other non-linear function is the oil production potential $q^r_{o,pot}$, which is a two-dimensional function of both the cumulative oil production N^r_p, and the field potential factor fn^r.

4.1. Problem reformulation

The non-linear functions G^r_p, W^r_p and $q^r_{o,pot}$ will be sampled from simulators and replaced with PWL approximations built from the sampled data. The simulated data are the outputs of the integrated reservoir-production model. The following functions will be approximated with PWL models:

$$G^r_p(t) = f_G(N^r_p(t)) \quad (35)$$

$$W^r_p(t) = f_W(N^r_p(t)) \quad (36)$$

$$q^r_{o,pot}(t) = fn^r(t) \times f_q(N^r_p(t)) \quad (37)$$

Further, notice that the multiplication of continuous variables in Eq. (37) yields a nonlinear constraint. To circumvent such nonlinearities, we approximate this multiplication also with the use of PWL functions.

The equations used in the multiplication linearization are the following:

$$q^r_{o,pot}(t) = fn^r(t) \times f_q(N^r_p(t)) \quad (38)$$

Notice that the production potential equation is presented twice, both in Eq. (37) and in Eq. (38). The reason for that is the presence of both the implicit function $f_q(N^r_p(t))$ and the multiplication term $fn^r(t) \times f_q(N^r_p(t))$.

Gas and water rate in time are computed by reformulating Eqs. (13) and (15) to Eqs. (39) and (40).

$$q^r_g(t) = G^r_p(t+1) - G^r_p(t) \quad (39)$$

$$q^r_w(t) = W^r_p(t+1) - W^r_p(t) \quad (40)$$

4.2. SOS2 Formulation

A continuous non-linear function $f(x) : \mathcal{D} \rightarrow \mathbb{R}^d$ with a compact domain $\mathcal{D}$ can be approximated with a set of linear functions, which are valid in a family of polytopes $\mathcal{P}$ with corresponding vertices $V(P)$, such that $\cup_{P \in \mathcal{P}} P = \mathcal{D}$, $\{\boldsymbol{m}_P\}_{P \in \mathcal{P}} \subseteq \mathbb{R}^d$, and $\{c_P\}_{P \subseteq \mathcal{P}}$, where:

$$f(x) = m'_p x + c_p,\ \forall x \in P, P \in \mathcal{P} \quad (41)$$

There are several different mathematical formulations for modelling PWL functions, see Vielma et al. (2010) for a review. A PWL formulation that has become popular for its efficiency and simplicity is known as Specially Ordered Sets of Type 2 (SOS2), see Beale and Tomlin (1970) and Beale (1980). The SOS2 model is based on a convex combination of weighting variables associated to breakpoints in the domain of the function of interest. The SOS2 formulation works by ensuring that at most two of such weighting variables can be nonzero simultaneously and, when that happens, they need to be consecutive for a given ordering of vertices in the domain. These constraints are typically imposed in the branch-and-bound algorithm by demand, and many off-the-shelf solvers have native support for SOS2 constraints.

4.2.1. PWL approximations using SOS2

We assume the function $f_G(N^r_p(t))$ is sampled over the domain $N^r_p(t)$ for a set of breakpoints $\mathcal{K}_G$ in $G^r_p(t)$ and the corresponding function values are denoted as $f^r_G(k)$. The PWL approximation of the non-linear function in Eq. (35) is formulated as follows:

$$\widetilde{G_p}^r(t) = \sum_{k \in \mathcal{K}_G} \eta^r_k(t) \cdot f^r_G(k) \quad (42)$$

$$N^r_p(t) = \sum_{k \in \mathcal{K}_G} \eta^r_k(t) \cdot N^r_p(k) \quad (43)$$

$$\sum_{k \in \mathcal{K}_G} \eta^r_k(t) = 1,\ \eta^r_k(t) \geq 0 \quad (44)$$

$$(\eta_k(t))_{k\in\mathcal{K}_G} \text{ is a SOS2} \quad (45)$$

where η_k^r are weighting variables used in the PWL approximation, and Eq. (45) are the SOS2 constraints which are imposed by the solver.

Analogously we sample the function $f_W(N_p^r(t))$ in a set of breakpoints $\mathcal{K}_W$ over the domain $N_p^r(t)$, and store the corresponding function values $f_W^r(k)$. The PWL linearization of Eq. (36) is then defined as follows:

$$\widetilde{W}_p^r(t) = \sum_{k\in\mathcal{K}_W} \sigma_k^r(t)\cdot f_W^r(k) \quad (46)$$

$$N_p^r(t) = \sum_{k\in\mathcal{K}_W} \sigma_k^r(t)\cdot N_p^r(k) \quad (47)$$

$$\sum_{k\in\mathcal{K}_W} \sigma_k^r(t) = 1,\ \sigma_k^r(t)\geq 0 \quad (48)$$

$$\left(\sigma_k^r(t)\right)_{k\in\mathcal{K}_W} \text{ is a SOS2} \quad (49)$$

where σ^r are the weighting variables used in the PWL approximation, and Eq. (49) are the corresponding SOS2 constraints.

Further, the function f_q is sampled in a set of breakpoints $\mathcal{K}_Q$ over the domain $N_p^r(t)$, and the corresponding function values $f_Q^r(k)$ are stored in table format. Since this function is used in the nonlinear multiplication in Eqs. (37) and (38), we linearize this relation using a 2-dimensional PWL function as follows:

$$q_{o,pot}^r(t) = \sum_{j\in\mathcal{K}_F}\sum_{k\in\mathcal{K}_Q} \Omega_{j,k,t}^r \cdot f_q^r(k) \quad (50)$$

$$fn^r = \sum_{j\in\mathcal{K}_F}\sum_{k\in\mathcal{K}_Q} \Omega_{j,k,t}^r \cdot fn^r(j) \quad (51)$$

$$\sum_{j\in\mathcal{K}_F}\sum_{k\in\mathcal{K}_Q} \Omega_{j,k,t}^r = 1,\ \Omega_{j,k,t}^r \geq 0 \quad (52)$$

$$\phi_j = \sum_{k\in\mathcal{K}_Q} \Omega_{j,k,t}^r,\ \forall j\in\mathcal{K}_F \quad (53)$$

$$\phi_k = \sum_{j\in\mathcal{K}_F} \Omega_{j,k,t}^r,\ \forall j\in\mathcal{K}_Q \quad (54)$$

$$\left(\phi_j\right)_{j\in\mathcal{K}_F} \text{ is a SOS2} \quad (55)$$

$$(\phi_k)_{k\in\mathcal{K}_Q} \text{ is a SOS2} \quad (56)$$

where the function fn^r is sampled in a set of breakpoints $\mathcal{K}_F$, $\Omega_{j,k,t}^r$ are the weighting variables for the PWL approximation, and ϕ_j and ϕ_k are auxiliary variables which are required in the PWL multidimensional approximation using SOS2. Eqs. (55) and (56) are the SOS2 constraints which are implemented by the solver.

4.3. Logarithmic formulation

The PWL function (41) can be described with several formulations other than the SOS2 formulation. These formulations vary in the way they represent the polytopes $P\in\mathcal{P}$ in the domain and the function approximation itself. Although all the different PWL formulations are equivalent in terms of accuracy for the same domain partitioning $\mathcal{P}\subseteq\mathcal{D}$, they can vary significantly in terms of size and efficiency. One crucial aspect of such formulations is the number of additional variables and constraints required to construct the approximation. The SOS2 formulation does not require any additional variables and constraints, but its performance tends to degrade for multidimensional approximations with a large number of breakpoints (Silva and Camponogara, 2014; Vielma et al., 2010).

A formulation which has promising properties for modeling multidimensional functions with numerous breakpoints is the Logarithmic formulation, also known as Log, see Vielma and Nemhauser (2011). Log is a variation of the aggregated convex combination (CC) model Keha et al. (2004); Lee and Wilson (2001); Padberg (2000) that requires an additional number of binary variables and constraints that grow logarithmically with respect to the number of breakpoints. Because of the compactness of the resulting formulation and the strength of its linear relaxations, Log generally enables considerable improvements in terms of efficiency compared to other PWL formulations. As Log relies on a convex combination of breakpoints of the function domain, one weighting variable is assigned to each vertex $v\in\mathcal{V}(\mathcal{P})$ of the domain such that the point in the graph of the function is described through a convex combination of the function values at the vertices, i.e, $(\mathbf{x}, f(\mathbf{x})) = \sum_{\mathbf{v}\in\mathcal{V}(\mathcal{P})}\lambda_{\mathbf{v}}(\mathbf{v}, f(\mathbf{v}))$, $\{\lambda_{\mathbf{v}}\}_{\mathbf{v}\in\mathcal{V}(\mathcal{P})}\subset\mathbb{R}_+$ such that $\sum_{\mathbf{v}\in\mathcal{V}(\mathcal{P})}\lambda_{\mathbf{v}} = 1$.

Although both SOS2 and Log rely on a convex combination of breakpoints of the domain, they differ significantly on how they create the domain partitioning $\mathcal{P}\subseteq\mathcal{D}$. SOS2 selects a single active polytope $P\in\mathcal{P}$ using on-demand constraints imposed by the optimization solver directly in the branching algorithm (Beale, 1980). On the other hand, Log utilizes a logarithmic number of additional binary variables and constraints to create a branching scheme that will select the active polytope within the domain. In other words, Log relies on an injective function $B:\mathcal{P}\to\{0,1\}^{\lceil\log_2|\mathcal{P}|\rceil}$ such that $B(P)=\mathbf{y}$ to map each polytope $P\in\mathcal{P}$ with a binary vector $\mathbf{y}\in\{0,1\}^{\lceil\log_2|\mathcal{P}|\rceil}$. The only requirement for the function B is that it must be compatible with SOS2 constraints, i.e., the non-zero λ variables need to be associated with the vertices of at least one polytope P of $\mathcal{P}$:

$$\exists P\in\mathcal{P} \text{ such that } \{\mathbf{v}\in\mathcal{V}(\mathcal{P}):\lambda_{\mathbf{v}}>0\}\subseteq V(P) \quad (57)$$

whereas the other λ variables that lie outside the active polytope P are equal to zero. For a 2D illustrative example of Log, see (Silva et al., 2012).

In Vielma and Nemhauser (2011) a branching scheme for Log is proposed for a valid injective function B. The proposed scheme generates a domain partitioning $\mathcal{P}$ that is topologically equivalent to a triangulation known as J1 or "Union Jack". This domain partitioning is created through a set of additional constraints, which consists of two phases. The first phase constrains the active polytope to a single hypercube using disjunctive sets. Then, in a second stage, certain vertices within the selected hypercube are disabled such that the convex combination is restricted to a single simplex.

The implementation of the branching scheme proposed by Vielma and Nemhauser (2011) requires new concepts and definitions. Let $\mathcal{S}_e=\{s_0,\ldots,s_n\}$ be the set of ordered breakpoints on the coordinate e, and $\mathcal{I}_e:=\{[s_0,s_1],\ldots,[s_{n-1},s_n]\}$ be the intervals containing pairs of consecutive breakpoints. Let $\mathcal{I}_e(s):=\{\mathcal{I}\in\mathcal{I}_e : s\in\mathcal{I}\}$ be a set of the intervals containing the breakpoint s, and $\Phi_e([s_i,s_{i+1}])=i+1$ be the index of an interval $[s_i,s_{i+1}]\in\mathcal{I}_e$. We define the function $B:\{1,\ldots,|\mathcal{I}_e|\}\to\{0,1\}^{\lceil\log_2(|\mathcal{I}_e|)\rceil}$ to be a mapping between the interval indices and a binary code according to the Gray code property, meaning that $B(i)$ and $B(i+1)$ must differ by only one bit. The vertices of the domain is $\mathcal{V}(\mathcal{P})=S_1\times\cdots\times S_d$ and d is the dimension.

The first phase of the branching scheme uses the sets $J_{e,B,l}^+ := \{s\in\mathcal{S}_e : B(\Phi_e(\mathcal{I}))_l = 1, \forall\mathcal{I}\in\mathcal{I}_e(s)\}$ and $J_{e,B,l}^0 := \{s\in\mathcal{S}_e : B(\Phi_e(\mathcal{I}))_l = 0, \forall\mathcal{I}\in\mathcal{I}_e(s)\}$. The constraints which implement the first phase of the Log branching scheme are defined as follows:

$$\sum_{\mathbf{v}\in\mathcal{V}^{+}_{e,B,l}} \lambda_{\mathbf{v}} \le x_{e,l},\ \forall e \in \{1,\dots,n\},\ l \in \{1,\dots,\lceil \log_2(|\mathcal{I}_e|)\rceil\} \quad (58a)$$

$$\sum_{\mathbf{v}\in\mathcal{V}^{0}_{e,B,l}} \lambda_{\mathbf{v}} \le x_{e,l},\ \forall e \in \{1,\dots,n\},\ l \in \{1,\dots,\lceil \log_2(|\mathcal{I}_e|)\rceil\} \quad (58b)$$

$$x_{e,l} \in \{0,1\},\ \forall e \in \{1,\dots,n\},\ l \in \{1,\dots,\lceil \log_2(|\mathcal{I}_e|)\rceil\} \quad (58c)$$

where $\mathcal{V}^{+}_{e,B,l} := \{\mathbf{v} \in \mathcal{V}(\mathcal{P}) : \mathbf{v}_e \in J^{+}_{e,B,l}\}$ and $\mathcal{V}^{0}_{e,B,l} := \{\mathbf{v} \in \mathcal{V}(\mathcal{P}) : \mathbf{v}_e \in J^{0}_{e,B,l}\}$. The sets $\mathcal{V}^{+}_{e,B,l}$ and $\mathcal{V}^{0}_{e,B,l}$ create the partitioning $\mathcal{P}$ in each coordinate e of the domain, and the intersection of the partitioning in all coordinates will constrain the domain to a single active hypercube.

The second phase selects a simplex of the hypercube obtained in phase one using the sets $\mathcal{L}_{r,s} = \{\mathbf{v} \in \mathcal{V}(\mathcal{P}) : \mathbf{v}_r \text{ is even and } \mathbf{v}_s \text{ is odd}\}$ and $\mathcal{R}_{r,s} = \{\mathbf{v} \in \mathcal{V}(\mathcal{P}) : \mathbf{v}_r \text{ is odd and } \mathbf{v}_s \text{ is even}\}, \forall r,s \in D = \{1,\dots,d\}$, such that $r < s$. The second branching phase can be implemented in Log with the following constraints:

$$\sum_{\mathbf{v}\in\mathcal{L}_{r,s}} \lambda_{\mathbf{v}} \le y_{r,s},\ \forall (r,s) \in \Gamma \quad (59a)$$

$$\sum_{\mathbf{v}\in\mathcal{R}_{r,s}} \lambda_{\mathbf{v}} \le 1 - y_{r,s},\ \forall (r,s) \in \Gamma \quad (59b)$$

$$y_{r,s} \in \{0,1\},\ \forall (r,s) \in \Gamma \quad (59c)$$

where $\Gamma := \{(r,s) \in \{1,\dots,d\} \times \{1,\dots,d\} : r < s\}$ is the set of index pairs indicating which weighting variables are to be disabled in the convex combination. The sets $\mathcal{L}_{r,s} := \{\mathbf{v} \in \mathcal{V} : \mathbf{v}_r \text{ is even and } \mathbf{v}_s \text{ is odd}\}$ and $\mathcal{R}_{r,s} := \{\mathbf{v} \in \mathcal{V} : \mathbf{v}_r \text{ is odd and } \mathbf{v}_s \text{ is even}\}$ create the partitioning responsible for scoping the active polytope to a simplex within the selected hypercube in phase 1.

4.3.1. PWL approximations using log

Based on the Log model we propose PWL approximations for the non-linear functions (35), (36), and (37). The Log PWL approximation of $G^r_p(t)$ defined in Eq. (35) is formulated as follows:

$$\widetilde{G_p}^r(t) = \sum_{k\in\mathcal{K}_G} \eta^r_k(t) \cdot f^r_G(k) \quad (60a)$$

$$N^r_p(t) = \sum_{k\in\mathcal{K}_G} \eta^r_k(t) \cdot N^r_p(k) \quad (60b)$$

$$\sum_{k\in\mathcal{K}_G} \eta^r_k(t) = 1,\ \eta^r_k(t) \ge 0 \quad (60c)$$

$$\sum_{k\in\mathcal{K}^{+}_{G,l}} \eta^r_k(t) \le x^{Gp}_l,\ l \in \{1,\dots,\lceil \log_2(|\mathcal{I}_{Gp}|)\rceil\} \quad (60d)$$

$$\sum_{k\in\mathcal{K}^{0}_{G,l}} \eta^r_k(t) \le 1 - x^{Gp}_l,\ l \in \{1,\dots,\lceil \log_2(|\mathcal{I}_{Gp}|)\rceil\} \quad (60e)$$

$$x^{Gp}_l \in \{0,1\},\ l \in \{1,\dots,\lceil \log_2(|\mathcal{I}_{Gp}|)\rceil\} \quad (60f)$$

where $\mathcal{K}^{+}_{G,l}$ and $\mathcal{K}^{0}_{G,l}$ are the first-phase branching sets for the set of ordered breakpoints $\mathcal{K}_G$, and $\mathcal{I}_{Gp}$ is the set of intervals containing the ordered pair of breakpoints in $\mathcal{K}^{0}_{G,l}$. These sets are defined analogously to the sets used in the first branching phase formulated with Eqs. (58a), (58b), and (60f).

Next, we formulate an approximation using Log for the function $f_W(N^r_p(t))$ defined in Eq. (36) with the following set of equations:

$$\widetilde{W_p}^r(t) = \sum_{k\in\mathcal{K}_W} \sigma^r_k(t) \cdot f^r_W(k) \quad (61)$$

$$N^r_p(t) = \sum_{k\in\mathcal{K}_W} \sigma^r_k(t) \cdot N^r_p(k) \quad (62)$$

$$\sum_{k\in\mathcal{K}_W} \sigma^r_k(t) = 1,\ \sigma^r_k(t) \ge 0 \quad (63)$$

$$\sum_{k\in\mathcal{K}^{+}_{W,l}} \sigma^r_k(t) \le x^{Wp}_l,\ l \in \{1,\dots,\lceil \log_2(|\mathcal{I}_{Wp}|)\rceil\} \quad (64)$$

$$\sum_{k\in\mathcal{K}^{0}_{W,l}} \sigma^r_k(t) \le 1 - x^{Wp}_l,\ l \in \{1,\dots,\lceil \log_2(|\mathcal{I}_{Wp}|)\rceil\} \quad (65)$$

$$x^{Wp}_l \in \{0,1\},\ l \in \{1,\dots,\lceil \log_2(|\mathcal{I}_{Wp}|)\rceil\} \quad (66)$$

with $\mathcal{K}^{+}_{W,l}$ and $\mathcal{K}^{0}_{W,l}$ being the first-phase branching sets, and $\mathcal{I}_{Wp}$ the set of intervals containing the ordered pair of breakpoints of $\mathcal{K}_W$. Notice that the Log approximations of both (35) and (36) use only the first branching phase. This is because the function domains are unidimensional, thus the active polytopes will be an interval belonging to $\mathcal{I}_{Gp}$ and $\mathcal{I}_{Wp}$.

The last function to be approximated with Log is the production potential f_q. As this function is present in a nonlinear multiplication of variables in Eqs. (37) and (38), we approximate these relations with a two-dimensional PWL approximation using Log as follows:

$$q^r_{o,pot}(t) = \sum_{j\in\mathcal{K}_F} \sum_{k\in\mathcal{K}_Q} \Omega^r_{j,k,t} \cdot f^r_q(k) \quad (67)$$

$$fn^r = \sum_{j\in\mathcal{K}_F} \sum_{k\in\mathcal{K}_Q} \Omega^r_{j,k,t} \cdot fn^r(j) \quad (68)$$

$$\sum_{j\in\mathcal{K}^{+}_{F,l}} \sum_{k\in\mathcal{K}_Q} \Omega^r_{j,k,t} \le x^{F,r}_{t,l},\ l \in \{1,\dots,\lceil \log_2(|\mathcal{I}_F|)\rceil\} \quad (69)$$

$$\sum_{j\in\mathcal{K}^{0}_{F,l}} \sum_{k\in\mathcal{K}_Q} \Omega^r_{j,k,t} \le 1 - x^{F,r}_{t,l},\ l \in \{1,\dots,\lceil \log_2(|\mathcal{I}_F|)\rceil\} \quad (70)$$

$$x^{F,r}_{t,l} \in \{0,1\},\ l \in \{1,\dots,\lceil \log_2(|\mathcal{I}_F|)\rceil\} \quad (71)$$

$$\sum_{j\in\mathcal{K}_F} \sum_{k\in\mathcal{K}^{+}_{Q,l}} \Omega^r_{j,k,t} \le x^{Q,r}_{t,l},\ l \in \{1,\dots,\lceil \log_2(|\mathcal{I}_Q|)\rceil\} \quad (72)$$

$$\sum_{j\in\mathcal{K}_F} \sum_{k\in\mathcal{K}^{0}_{F,l}} \Omega^r_{j,k,t} \le 1 - x^{Q,r}_{t,l},\ l \in \{1,\dots,\lceil \log_2(|\mathcal{I}_Q|)\rceil\} \quad (73)$$

$$x^{Q,r}_{t,l} \in \{0,1\},\ l \in \{1,\dots,\lceil \log_2(|\mathcal{I}_Q|)\rceil\} \quad (74)$$

$$\sum_{(j,k)\in\mathcal{L}_{j,k}} \Omega^r_{j,k,t} \le y^r_{j,k,t},\ \forall (j,k) \in \Gamma^r_t \quad (75)$$

$$\sum_{(j,k)\in\mathcal{R}_{j,k}} \Omega^r_{j,k,t} \le 1 - y^r_{j,k,t},\ \forall (j,k) \in \Gamma^r_t \quad (76)$$

Table 6
Parameters of the reservoir and network model.

Reservoir & Well		
Parameter	Reservoir-1	Reservoir-2
Reservoir Pressure, (bara)	195	243
Reservoir Temperature, (^oC)	70	90
Oil in place, (MSm^3)	56.25	39.25
Solution gas-oil ratio, (Sm^3/Sm^3)	115	150
Initial water saturation, (fraction)	0.05	0.05
Number of wells	6	3
Productivity index, ($Sm^3/d/bar$)	1500	500
Tubing size, (inch)	5.5	5.5

Surface Network	
System type	Production
Seabed Temperature, (^oC)	4
Pipeline diameter, (inch)	6 - 10
Pipeline Length, (Km)	17.5
Separator Pressure, (bara)	20

$$y^r_{j,k,t} \in \{0,1\},\ \forall (j,k) \in \Gamma^r_t \tag{77}$$

where Eqs. (69)–(71) implement the first branching phase of the Log for the set $\mathcal{K}_{\mathrm{F}}$, whereas Eqs. (72)–(74) are responsible for the first phase branching for set $\mathcal{K}_{\mathrm{Q}}$. The second phase branching scheme is implemented by Eqs. (75)–(77). Notice that the sets $\mathcal{K}^+_{\mathrm{F,l}}$, $\mathcal{K}^0_{\mathrm{F,l}}$, $\mathcal{K}^+_{\mathrm{Q,l}}$, $\mathcal{K}^0_{\mathrm{F,l}}$, $\mathcal{I}_{\mathrm{F}}$, and $\mathcal{I}_{\mathrm{Q}}$ are defined analogously to the definitions of the first phase branching in Eqs. (58a)–(58c). The sets $\mathcal{L}_{j,k}$, $\mathcal{R}_{j,k}$, and Γ^r_t on its turn are defined analogously to the definitions used in the second phase branching scheme denoted by Eqs. (59a)–(59c).

5. Simulations

In this section, we present a computational analysis assessing the accuracy and performance of the proposed formulations, both SOS2 and Logarithmic, in field development optimization problems, and a case study of real field producing from 2 reservoirs for 20 years. The production potential curves for the different well combinations were generated using the commercial software *Petroleum Expert IPM* Experts (2008) with a material balance model for the reservoir using MBAL coupled to a network model represented with GAP. Reservoir and network parameters of the coupled simulations are listed in Table 6. The simulation results of fractional factors with different well combinations are provided in the Appendix. The computational analysis compares the performance and approximation accuracy of the PWL models for field development problems of different complexities in order to demonstrate how such models scale with the number of variables and constraints. The case study aims to demonstrate the effectiveness of the Log model in large-scale field development problems both in terms of efficiency and in terms of approximation accuracy of the final results.

5.1. Performance and accuracy study of PWL models

In order to assess the effectiveness of the SOS2 and the Log models applied to field development optimization, we perform a computational study with some representative problems and compare the performance of both models in terms of accuracy and efficiency on a ThinkPad of Intel(R) Core(TM) i7-8565U CPU @ 1.80 Hz 1.99 GHz 64 bytes. It is expected the objective function to exhibit some differences when using Log (simplices) or SOS2 (hypercubes) for the PWL approximation (In Section 4) but the values should still be comparable. Also, when the production horizon of the field is increased, the optimization problem becomes harder, and the optimal solution with 0% of dual gap (optimality certificate) might not be obtained within a reasonable time.

Table 7
Performance Comparison of SOS2 vs. Log (with 5% dual gap stopping criteria).

Model	No.	Production horizon	CPU	Gap	NPV
SOS2/LOG		year	second	%	USD
SOS2	1	3	1.35	5.00	3,291,430,000
	2	4	43.80	5.00	3,666,370,000
	3	5	115.98	5.00	4,209,920,000
	4	6	1,540.57	5.00	4,617,210,000
	5	7	2,760.23	5.00	4,953,380,000
	6	8	20,851.39	5.00	5,220,280,000
LOG	7	3	2.13	4.98	3,291,430,000
	8	4	4.95	4.87	3,666,370,000
	9	5	9.94	4.76	4,209,920,000
	10	6	11.66	4.90	4,612,050,000
	11	7	14.45	4.97	4,953,380,000
	12	8	53.35	4.91	5,234,550,000

Table 8
Performance Comparison of SOS2 vs. Log (with 500 seconds stopping criteria)

Model	No.	Production horizon	CPU	Gap	NPV
SOS2/LOG		year	second	%	USD
SOS2	1	3	25.26	0.0	3,291,430,000*
	2	4	176.71	0.0	3,666,370,000*
	3	5	500.00	2.04	4,209,920,000*
	4	6	500.00	5.30	4,617,210,000*
	5	7	500.00	6.59	4,914,570,000**
	6	8	500.00	7.21	5,177,310,000**
LOG	7	3	3.88	0.0	3,291,430,000*
	8	4	7.15	0.0	3,666,370,000*
	9	5	15.32	0.0	4,209,920,000*
	10	6	52.15	0.0	4,617,210,000*
	11	7	67.14	0.0	4,953,800,000**
	12	8	98.81	0.0	5,238,500,000**

Two batches of simulations were performed to test the performance of the PWL formulations. In the first we set a dual gap of $< 5\%$ as the stopping criteria for the algorithms, and compare the CPU time for the optimization with both PWL models. For the second set of simulations, a time limit of 500 seconds is set to be the stopping criteria, and the dual gap of the final solution obtained by the different PWL models are compared.

The results of the tests using a dual gap value of 5% as stopping criteria are shown in Table 7. The table presents the optimization's running time (in CPU seconds), the value of the objective function (NPV) and the value of the gap for six values of production horizons and when using the SOS2 and Log model. The production horizon was varied between 3 and 8 years.

When using the SOS2 model to solve the 3 years production horizon problem, an optimal solution is obtained in 1.35 CPU seconds, while it takes 2.13 CPU seconds using the Log model, which is a small difference. However, for longer production horizons the difference becomes substantial, i.e., for 8 years the SOS2 model takes 20,851.39 CPU seconds (ca. 348 minutes) and the Log model, takes 53.35 CPU seconds (ca. 1 minute). The comparison clearly indicates the Log model is more efficient to solve the optimization problem.

In the second batch of simulations, a stopping criteria of 500 CPU seconds was used. The computational results are presented in Table 8. The objective values marked with a single asterisk (*) indicate that simulations using SOS2 and Log models computed the same optimum. However for some cases the results using the SOS2 model have a gap greater than zero at the end of the run. When using the SOS2 model, a longer production horizon positively correlated to a higher value of the gap (e.g., 2.04% gap in 5 years lifetime and 5.30% gap in 6 years lifetime).

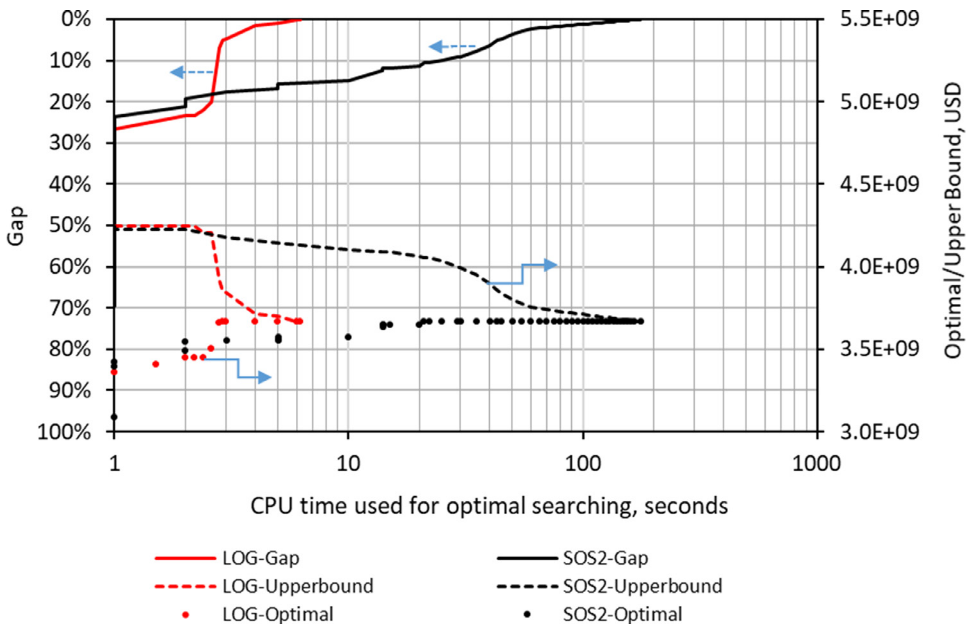


Fig. 4. Optimal searching process using SOS2 and Log algorithm.

Objective values marked with two asterisks (**) represent cases where large differences were detected in the objective function at the end of the run when using the SOS2 and Log models. Simulations performed with the Log model achieved higher values of the objective function with zero gaps.

Based on the results presented in Table 7 and Table 8, it is possible to conclude that the Log model enables solving the field development problem more efficiently than the SOS2 model.

Fig. 4 shows the value of the gap and the objective function versus CPU time when using SOS2 and Log models and using a production horizon of 4 years. The formulation using SOS2 models requires around 176.71 CPU seconds to find a solution with 0% of dual gap. In contrast, the formulation using Log models required 7.15 CPU seconds only to reach a dual gap value of 0%, i.e., 24 times faster. Even though the optimal solution with the SOS2 model was obtained within around 20 CPU seconds, close to the time spent by Log (7.15 CPU seconds), it takes a long time to prove optimality of the solution, i.e, to close the dual gap upper-bound (the black dashed line).

5.2. Case study: real-world multi-reservoir field development optimization

The Log model was applied to the study case in offshore oil field development presented in Section 2. This case study aims to test the effectiveness of the logarithmic formulation in a realistic, large-scale model. Some parameters and information about the study case are provided next.

The field has 2 independent reservoirs and is designed to produce for 20 years. After performing geology and petroleum engineering studies, 9 wells with pre-specified paths and placement positions are considered as drilling candidates, of which $(w1, w2, w3, w4, w5, w6)$ are placed in Reservoir 1 and $(w7, w8, w9)$ placed in Reservoir 2. A black-box simulation model was built considering the reservoir and the production facilities. Production potential curves were generated using this model for all possible well combinations.

The decision variables of the optimization problem are the well allocation and drilling schedule of the field for its lifetime such that its NPV is maximized.

Table 9
Constraints information.

Constraints	Value:
Lifetime:	20 *years*
Oil price P_o:	60 *USD/bbl*
Gas price P_g:	2 *USD/MMBTU*
Maximum drilling capacity N_w^{Dmax}:	3 *wells/year*
Maximum oil processing capacity q_o^{max}:	3650×10^3 $m^3/year$
Maximum Gas processing capacity q_g^{max}:	2.2×10^9 $m^3/year$
Capital return period N_D:	4 *years*

Table 10
Formulation size & computational performance.

Variables	
Binary variables:	770
Integer variables:	60
Linear variables:	7960

Constraints	
Equality constraints:	782
Inequality constraints:	3469
Range constraints:	19

Solving information	
Solver:	Gurobi
CPU time:	2161.75 seconds
GAP:	0%

Three wells from Reservoir 1 are considered to be pre-drilled before production in the first year. The number of wells that can be drilled each year should be less or equal to 3. All the other constraints are presented in Table 9. The objective function is the maximization of the NPV, accounting for the revenue obtained from oil and gas sales, and the cost from drilling the wells, and with the facilities construction and operations costs which in function of the field production rates (oil, gas and water).

Table 10 presents a summary of number and type of variables and constraints employed in the model and the solver details and stopping criteria. The overall optimization model consists of 8790 variables and 4270 constraints. The problem is formulated

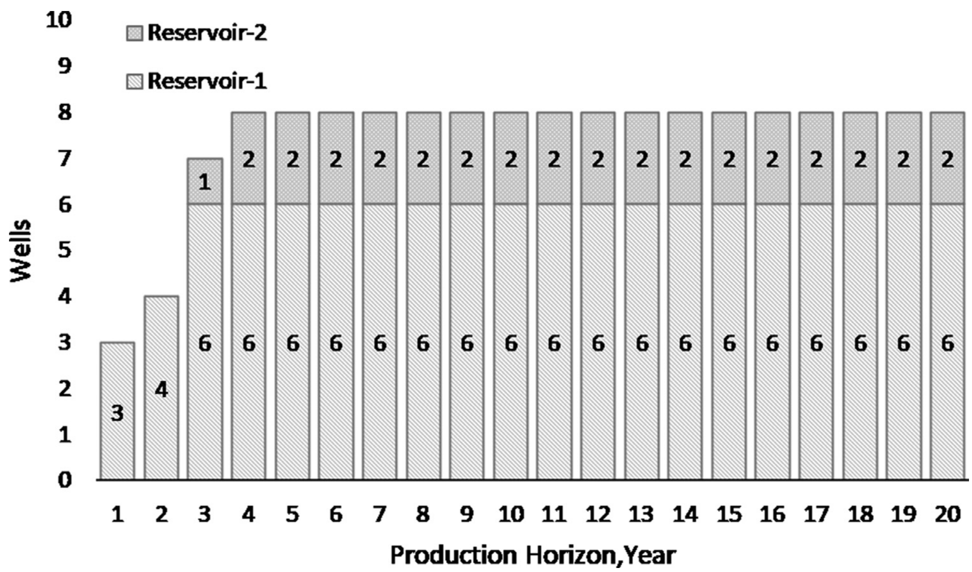


Fig. 5. Optimal drilling schedule.

Table 11
Drilling well sequence.

Year	Well in production	Reservoir-1	Reservoir-2
1	3	$w1, w4, w5$	/
2	4	$w1, w3, w4, w5$	/
3	7	$w1, w2, w3, w4, w5, w6$	$w9$
4	8	$w1, w2, w3, w4, w5, w6$	$w8, w9$
...	...	...	...
20	8	$w1, w2, w3, w4, w5, w6$	$w8, w9$

using AMPL (Fourer et al., 2003) and solved with Gurobi Optimization (2020). The CPU time used to run the optimization problem and obtain the optimal solution (with 0% of dual gap) was 2161.75 seconds.

Fig. 5 and Table 11 show the optimal drilling schedule. As it can be seen, 3 specified wells ($w1, w4, w5$) from Reservoir 1 are set to start producing from the first year. A new well ($w3$) from Reservoir 1 is planned to start producing from the second year. In order to maintain the production plateau of the field, 3 new wells are planned to start producing from the third year, 2 wells ($w2, w6$) from Reservoir 1 and 1 well ($w9$) from Reservoir 2. In the fourth year, another well ($w8$) from Reservoir 2 is scheduled to start producing. In total, 8 wells are planned to be drilled and to produce from this field from the 9 available candidates. The optimization model determines the number of wells to be drilled in each year and which wells are to drill. The optimal solution consisted of 8 wells for the offshore field development, where the candidate well ($w7$) from Reservoir 2 is decided not to be drilled. The optimal well schedule honors all drilling-related constraints for the given parameters listed in Table 9.

Fig. 6 depicts the field yearly oil rates obtained by the optimization. The dashed lines represent the production potential and the solid lines represent actual oil yearly rates. It can be seen that this field has a production plateau of 4 years (black solid line). Most of the field production comes from Reservoir 1 (green solid line), although its production drops below that of Reservoir 2(red solid line) in the fourth year. All recoverable reserves in Reservoir 2 are expected to be produced after 17 years, whereas the production is expected to last 19 years in Reservoir-1.

Fig. 7 shows the optimum production rate and potential curves in function of cumulative oil production. The optimum oil production rate (solid lines) is below the maximum feasible value (production potential, in dashed lines). The production potential depends on the drilling schedule presented in Table 11.

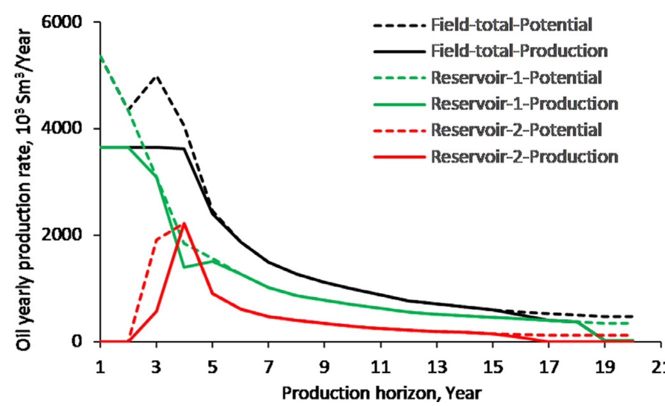


Fig. 6. Production potential Vs. optimal production.

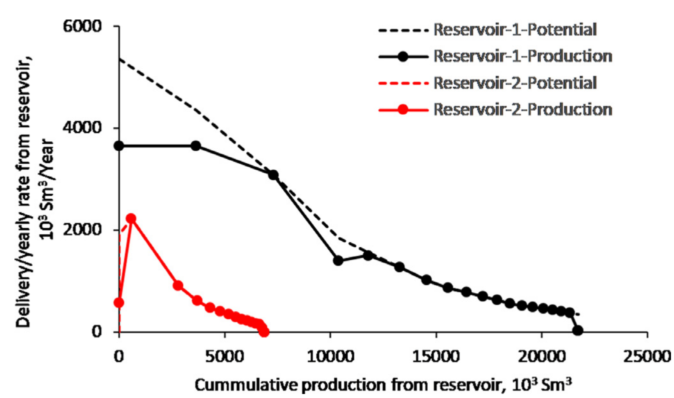


Fig. 7. Production potential Vs. optimal production.

6. Conclusions

We proposed a formulation using mathematical programming for field development optimization in early-phase of an offshore hydrocarbon field. The optimization consists in determining the drilling and production schedules that maximize the project value for a multi-reservoir field.

The model is formulated in a flexible manner such that it can be extended and customized to other field development problems. The main contribution of this work is two-fold. First we proposed a novel way to represent the drilling schedule and the well combination selection from all possible well permutations in the field. As the efficiency of the optimization is significantly impacted by the number of binary variables, we formulated the well combination selection with a set of constraints that require only a reduced number of binary variables, equal to the number of wells. This allows to scale the problem to larger fields with numerous wells.

Secondly we propose the use of a Logarithmic model to transform the field development MINLP problem into a MILP formulation. We demonstrate through a real-world case study that the Logarithmic formulation is substantially more efficient than the traditional SOS2 models, specially when the production horizon is more than just a few years. The logarithmic formulation also allows to reach lower dual gap values in a shorter time. Based on simulation analysis, we have the following specific conclusions:

- Both SOS2 and Log models have been applied to solving field development problems and the results show significant improvements in computational efficiency when using the Log model.
- More computational time is required to find optimal solutions when increasing the production horizon (field lifetime). The required computational time increased dramatically in the SOS2 model when compared to the Log model.
- The Log algorithm takes less than 1 h to find the optimal solution to a real field planning problem with a production horizon of 20 years.

The authors believe that the proposed optimization model using production potential curves is appropriate for field development of early phases when limited data is available, and reservoir models are highly uncertain, under construction, or unavailable. The model is suitable to run extensive analyses to evaluate uncertainty with a reduced computational budget. However, this approach may not be appropriate for later stages of the field development process when more complex models are used, such as compartmentalized, highly heterogeneous reservoirs, or capturing well placement issues.

Moreover, in the proposed optimization model, it is impossible to enforce constraints on the reservoir pressure and individual well rates as these variables are not tracked. This is an interesting aspect to be investigated in future works, as in real-world fields there are often constraints that need to be imposed on specific wells due to physical limitations or operational issues.

Further, we have used a time-step of one year based on previous works and to ensure compatibility with drilling constraints in our work. However, we believe it is important to evaluate the effect of the time step on the optimization results output by the model, and we suggest this as future work.

Another assumption made in this work is that the wells can produce at their potential rate. However, in some cases, this might not be possible due to technical constraints (e.g., sand production). As a workaround, it is possible to add additional constraints to the formulation to avoid field production reaching undesirably high levels.

We believe the proposed methodology can be a valuable decision-support tool for field planners, capturing first-order magnitude effects and output variability considering several uncertain parameters. Possible extensions of this work are to perform the planning and scheduling of multiple fields in the same area, determine the optimal location of subsea and topside facilities, and study staged developments, where decisions are taken sequentially, and models are updated with new information.

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.

CRediT authorship contribution statement

Guowen Lei: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing - original draft. **Thiago Lima Silva:** Conceptualization, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Writing - original draft. **Milan Stanko:** Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Writing - review & editing.

Acknowledgements

This research is a part of BRU21 NTNU Research and Innovation Program on Digital and Automation Solutions for the Oil and Gas Industry (www.ntnu.edu/bru21) and supported by Lundin Energy Norway. We thank the reviewers for their valuable inputs and contributions to the manuscript.

Appendix A

Table 12
Fractional Factors of Different Well Combinations.

Reservoir-1			
Well combinations	f_n^1	Well combinations	f_n^1
w1	0.166	w2	0.249
w3	0.195	w4	0.293
w5	0.198	w6	0.113
w1,w2	0.332	w1,w3	0.357
w1,w4	0.413	w1,w5	0.365
w1,w6	0.279	w2,w3	0.436
w2,w4	0.467	w2,w5	0.448
w2,w6	0.362	w3,w4	0.477
w3,w5	0.393	w3,w6	0.308
w4,w5	0.492	w4,w6	0.406
w5,w6	0.304	w1,w2,w3	0.516
w1,w2,w4	0.524	w1,w2,w5	0.530
w1,w2,w6	0.445	w1,w3,w4	0.591
w1,w3,w5	0.556	w1,w3,w6	0.470
w1,w4,w5	0.611	w1,w4,w6	0.526
w1,w5,w6	0.470	w2,w3,w4	0.642
w2,w3,w5	0.635	w2,w3,w6	0.549
w2,w4,w5	0.665	w2,w4,w6	0.580
w2,w5,w6	0.553	w3,w4,w5	0.676
w3,w4,w6	0.590	w3,w5,w6	0.499
w4,w5,w6	0.597	w3,w4,w5,w6	0.781
w2,w4,w5,w6	0.771	w2,w3,w5,w6	0.740
w2,w3,w4,w6	0.755	w2,w3,w4,w5	0.841
w1,w4,w5,w6	0.717	w1,w3,w5,w6	0.661
w1,w3,w4,w6	0.704	w1,w3,w4,w5	0.790
w1,w2,w5,w6	0.636	w1,w2,w4,w6	0.637
w1,w2,w4,w5	0.722	w1,w2,w3,w6	0.629
w1,w2,w3,w5	0.714	w1,w2,w3,w4	0.696
w1,w2,w3,w4,w5	0.895	w1,w2,w3,w4,w6	0.809
w1,w2,w3,w5,w6	0.820	w1,w2,w4,w5,w6	0.828
w1,w3,w4,w5,w6	0.895	w2,w3,w4,w5,w6	0.946
w1,w2,w3,w4,w5,w6	1.000		

Reservoir-2			
Well combinations	f_n^2	Well combinations	f_n^2
w7	0.527	w8	0.606
w9	0.644	w7,w8	0.807
w7,w9	0.878	w8,w9	0.907
w7,w8,w9	1.000		

References

Alkindira, S., 2020. Numerical Optimization Applied to Field Development. NTNU.
Angga, I., 2019. Automated Decision Support Methodologies for Field Development: The Safari Field Case. NTNU.
Aronofsky, J., Williams, A., 1962. The use of linear programming and mathematical models in under-ground oil production. Manage Sci 8 (4), 394–407.
Attra, H., Wise, W., Black, W., et al., 1961. Application of optimizing techniques for studying field producing operations. Journal of Petroleum Technology 13 (01), 82–86. doi:10.2118/1338-G-PA.
Beale, E.M.L., 1980. Branch and bound methods for numerical optimization of non–convex functions. Comput Stat 80, 11–20.
Beale, E.M.L., Tomlin, J.A., 1970. Special facilities in a general mathematical programming system for non-convex problems using ordered sets of variables. In: Proceedings of the Fifth International Conference on Operations Research, pp. 447–454.
Brito, B., Finardi, E., Takigawa, F., 2020. Mixed-integer nonseparable piecewise linear models for the hydropower production function in the unit commitment problem. Electr. Power Syst. Res. 182, 106234.
Carvalho, M., Pinto, J., 2006. An milp model and solution technique for the planning of infrastructure in offshore oilfields. Journal of Petroleum Science and Engineering 51 (1–2), 97–110.
Coats, B., Fleming, G., Watts, J., Rame, M., Shiralkar, G., et al., 2003. A generalized wellbore and surface facility model, fully coupled to a reservoir simulator. SPE Reservoir Simulation Symposium. Society of Petroleum Engineers.
Devine, M., Lesso, W., 1972. Models for the minimum cost development of offshore oil fields. Manage Sci 18 (8), B–378.
Durrer, E., Slater, G., 1977. Optimization of petroleum and natural gas productiona survey. Manage Sci 24 (1), 35–43.
Epelle, E.I., Gerogiorgis, D.I., 2019. Optimal rate allocation for production and injection wells in an oil and gas field for enhanced profitability. AIChE J. 65 (6), e16592.
Epelle, E.I., Gerogiorgis, D.I., 2020. A computational performance comparison of milp vs. minlp formulations for oil production optimisation. Computers & Chemical Engineering 106903.
Experts, P., 2008. IPM: Engineering software development. product brochure, dated.
Fourer, R., Gay, D. M., Kernighan, B. W., 2003. Ampl. a modeling language for mathematical programming. URL: http://www.ampl.com.
Frair, L., Devine, M., 1975. Economic optimization of offshore petroleum development. Manage Sci 21 (12), 1370–1379.
Garcia-Diaz, J.C., Startzman, R., Hogg, G.L., et al., 1996. A new methodology for minimizing investment in the development of offshore fields. SPE Production & Facilities 11 (01), 22–29.
Goel, V., Grossmann, I.E., 2004. A stochastic programming approach to planning of offshore gas field developments under uncertainty in reserves. Computers & Chemical Engineering 28 (8), 1409–1429.
Goel, V., Grossmann, I.E., El-Bakry, A.S., Mulkay, E.L., 2006. A novel branch and bound algorithm for optimal development of gas fields under uncertainty in reserves. Computers & Chemical Engineering 30 (6–7), 1076–1092.
González, D., Stanko, M., Hoffmann, A., 2019. Decision support method for early-phase design of offshore hydrocarbon fields using model-based optimization. J. Pet. Explor. Prod. Technol. 1–23.
Grimmett, T., Startzman, R., et al., 1987. Optimization of offshore field development to minimize investment. SPE Hydrocarbon Economics and Evaluation Symposium. Society of Petroleum Engineers.
Grossmann, I.E., Apap, R.M., Calfa, B.A., García-Herreros, P., Zhang, Q., 2016. Recent advances in mathematical programming techniques for the optimization of process systems under uncertainty. Computers & Chemical Engineering 91, 3–14.
Gunnerud, V., Foss, B., 2010. Oil production optimizationa piecewise linear model, solved with two decomposition strategies. Computers & Chemical Engineering 34 (11), 1803–1812.
Gunnerud, V., Foss, B.A., McKinnon, K., Nygreen, B., 2012. Oil production optimization solved by piecewise linearization in a branch & price framework. Computers & Operations Research 39 (11), 2469–2477.
Gupta, V., Grossmann, I.E., 2012. An efficient multiperiod minlp model for optimal planning of offshore oil and gas field infrastructure. Industrial & Engineering Chemistry Research 51 (19), 6823–6840.
Gupta, V., Grossmann, I.E., 2014. Multistage stochastic programming approach for offshore oilfield infrastructure planning under production sharing agreements and endogenous uncertainties. Journal of Petroleum Science and Engineering 124, 180–197.
Gupta, V., Grossmann, I.E., 2014. A new decomposition algorithm for multistage stochastic programs with endogenous uncertainties. Computers & Chemical Engineering 62, 62–79.
Gupta, V., Grossmann, I.E., 2017. Offshore oilfield development planning under uncertainty and fiscal considerations. Optimization and Engineering 18 (1), 3–33.
Gurobi Optimization, L., 2020. Gurobi optimizer reference manual. http://www.gurobi.com.
Hansen, P., de Luna Pedrosa Filho, E., Ribeiro, C.C., 1992. Location and sizing of offshore platforms for oil exploration. Eur J Oper Res 58 (2), 202–214.
Haugland, D., Hallefjord, Å., Asheim, H., 1988. Models for petroleum field exploitation. Eur J Oper Res 37 (1), 58–72.
Hepguler, G., Barua, S., Bard, W., et al., 1997. Integration of a field surface and production network with a reservoir simulator. SPE Computer Applications 9 (03), 88–92.
Hoffmann, A., Stanko, M., González, D., 2019. Optimized production profile using a coupled reservoir-network model. J. Pet. Explor. Prod. Technol. 9 (3), 2123–2137.
Humphries, T.D., Haynes, R.D., 2015. Joint optimization of well placement and control for nonconventional well types. Journal of Petroleum Science and Engineering 126, 242–253.
Huppler, J.D., et al., 1974. Scheduling gas field production for maximum profit. Society of Petroleum Engineers Journal 14 (03), 279–294.
Isebor, O.J., Echeverria Ciaurri, D., Durlofsky, L.J., et al., 2013. Generalized field development optimization using derivative-free procedures. SPE Reservoir Simulation Symposium. Society of Petroleum Engineers.
Iyer, R., Grossmann, I., Vasantharajan, S., Cullick, A., 1998. Optimal planning and scheduling of offshore oil field infrastructure investment and operations. Industrial & Engineering Chemistry Research 37 (4), 1380–1397.
Jahn, F., Cook, M., Graham, M., 2008. Hydrocarbon exploration and production. Elsevier.
Keha, A.B., de Farias Jr, I.R., Nemhauser, G.L., 2004. Models for representing piecewise linear cost functions. Operations Research Letters 32 (1), 44–48.
Khor, C.S., Elkamel, A., Shah, N., 2017. Optimization methods for petroleum fields development and production systems: a review. Optimization and Engineering 18 (4), 907–941.
Klose, A., Drexl, A., 2005. Facility location models for distribution system design. Eur J Oper Res 162 (1), 4–29.
Lee, A. S., Aronofsky, J. S., 1958. A linear programming model for scheduling crude oil production. Journal of Petroleum Technology 10 (07), 51–54.
Lee, J., Wilson, D., 2001. Polyhedral methods for piecewise-linear functions i: the lambda method. Discrete Appl. Math. 108 (3), 269–285.
Lin, X., Floudas, C.A., 2003. A novel continuous-time modeling and optimization framework for well platform planning problems. Optimization and Engineering 4 (1–2), 65–95.
McFarland, J.W., Lasdon, L., Loose, V., 1984. Development planning and management of petroleum reservoirs using tank models and nonlinear programming. Oper Res 32 (2), 270–289.
NPD, 2020. resource-report-exploration-2020-final.
Nygreen, B., Christiansen, M., Haugen, K., Bjørkvoll, T., Kristiansen, Ø., 1998. Modeling norwegian petroleum productionand transportation. Ann Oper Res 82, 251–268.
Padberg, M., 2000. Approximating separable nonlinear functions via mixed zero-one programs. Operations Research Letters 27 (1), 1–5.
Pavlov et al, A., 2020. BRU21 annual report 2019 NTNU research and innovation program in digital and automation solutions for the oil and gas industry. http://www.ntnu.edu/bru21.
Rahmawati, S.D., Whitson, C.H., Foss, B., Kuntadi, A., 2012. Integrated field operation and optimization. Journal of Petroleum Science and Engineering 81, 161–170.
Rodrigues, H., Prata, B.d.A., Bonates, T., 2016. Integrated optimization model for location and sizing of offshore platforms and location of oil wells. Journal of Petroleum Science and Engineering 145, 734–741.
Rosa, V.R., Camponogara, E., Ferreira Filho, V.J.M., 2018. Design optimization of oilfield subsea infrastructures with manifold placement and pipeline layout. Computers & Chemical Engineering 108, 163–178.
Rosenwald, G.W., Green, D.W., et al., 1974. A method for determining the optimum location of wells in a reservoir using mixed-integer programming. Society of Petroleum Engineers Journal 14 (01), 44–54.
Rowan, G., Warren, J., et al., 1967. Systems approach to reservoir engineering: optimum development planning. J. Can. Pet. Technol. 6 (03), 84–94.
Silva, T.L., Camponogara, E., 2014. A computational analysis of multidimensional piecewise-linear models with applications to oil production optimization. Eur J Oper Res 232 (3), 630–642.
Silva, T.L., Codas, A., Camponogara, E., 2012. A computational analysis of convex combination models for multidimensional piecewise-linear approximation in oil production optimization. IFAC Proceedings Volumes 45 (8), 292–298.
Silva, T.L., Codas, A., Stanko, M., Camponogara, E., Foss, B., et al., 2019. Network-constrained production optimization by means of multiple shooting. SPE Reservoir Evaluation & Engineering 22 (02), 709–733.
Stanko, M., 2021. Observations on and use of curves of current dimensionless potential versus recovery factor calculated from models of hydrocarbon production systems. Journal of Petroleum Science and Engineering 196, 108014.
Sullivan, J., 1988. The application of mathematical programming methods to oil and gas field development planning. Math Program 42 (1–3), 189–200.
Tarhan, B., Grossmann, I.E., Goel, V., 2009. Stochastic programming approach for the planning of offshore oil or gas field infrastructure under decision-dependent uncertainty. Industrial & Engineering Chemistry Research 48 (6), 3078–3097.
Tavallali, M.S., Karimi, I., Baxendale, D., 2016. Process systems engineering perspective on the planning and development of oil fields. AIChE J. 62 (8), 2586–2604.
Tavallali, M., Karimi, I., Teo, K., Baxendale, D., Ayatollahi, S., 2013. Optimal producer well placement and production planning in an oil reservoir. Computers & Chemical Engineering 55, 109–125.
Van Den Heever, S.A., Grossmann, I.E., 2000. An iterative aggregation/disaggregation approach for the solution of a mixed-integer nonlinear oilfield infrastructure planning model. Industrial & Engineering Chemistry Research 39 (6), 1955–1971.
van den Heever, S.A., Grossmann, I.E., Vasantharajan, S., Edwards, K., 2001. A lagrangean decomposition heuristic for the design and planning of offshore hydrocarbon field infrastructures with complex economic objectives. Industrial & Engineering Chemistry Research 40 (13), 2857–2875.

Vielma, J.P., Ahmed, S., Nemhauser, G., 2010. Mixed-integer models for nonseparable piecewise-linear optimization: unifying framework and extensions. Oper Res 58 (2), 303–315.
Vielma, J.P., Nemhauser, G.L., 2011. Modeling disjunctive constraints with a logarithmic number of binary variables and constraints. Math Program 128 (1–2), 49–72.
Wang, Y., Estefen, S.F., Lourenço, M.I., Hong, C., 2019. Optimal design and scheduling for offshore oil-field development. Computers & Chemical Engineering 123, 300–316.