Petroleum Science 19 (2022) 2431–2447

Contents lists available at ScienceDirect

Petroleum Science

journal homepage: www.keaipublishing.com/en/journals/petroleum-science

Original Paper

An optimization method for shale gas gathering system - Consideration of reliability enhancement under earthquake-related uncertainties

Yan Wu [a], Zi-Yuan Cui [a], Hai Lin [a], Yu-Fei Wang [a, *], Xiao Feng [b]

[a] School of Chemical Engineering and Environment, China University of Petroleum (Beijing), 18 Fuxue Road, Changping, Beijing, 102249, China
[b] School of Chemical Engineering and Technology, Xi'an Jiaotong University, 28 Xianning West Road, Xi'an, Shaanxi, 710049, China

ARTICLE INFO

Article history:
Received 11 October 2021
Received in revised form
29 December 2021
Accepted 29 December 2021
Available online 1 January 2022

Edited by Xiu-Qiu Peng

Keywords:
Shale gas
Gathering system
Network topology
Reliability
Uncertainty

ABSTRACT

Shale gas gathering systems are of great importance in gas fields for the efficient and reliable transportation of gas. In traditional design methods, pipeline layouts are optimized only from an economic point of view, where reliability is evaluated after optimization. However, reliability can be enhanced by spare pipelines so that both economic and reliable aspects are evaluated simultaneously. Based on the idea of enhancing reliability, this research proposes a methodology for optimizing the pipeline layout problem including well clustering, stations site selecting, and piping. Different topology arrangements and spare pipelines are investigated to enhance the production efficiency and reliability of the pipeline network under earthquake-related uncertainties. Reliability evaluation is converted into an economic one so that the objective of this work is to minimize the total annual cost. To solve such a complex problem, genetic algorithm, K-means algorithm, GeoSteiner algorithm, Kruskal algorithm, linear programming and Monte Carlo simulations are combined. A real-world case study illustrates the effectiveness of the proposed methodology and shows a 36.53% reduction in the total annual cost compared with the initial scheme.

1. Introduction

As a potential unconventional natural gas resource, shale gas has a long exploitation life and a long stable production cycle. With the continuous increase of energy demand, the status of shale gas in energy consumption has gradually increased. Shale gas reserves are abundant worldwide. According to the assessment of the Energy Information Administration (EIA) (EIA and DoE, 2013), technically recoverable shale gas resources can reach 2.21×10^{14} m^3. By 2040, the global daily natural gas production will rise from 9.7×10^9 m^3 in 2015 to nearly 1.57×10^{10} m^3, in which shale gas production will account for 30% of the global production (Wang, 2018). The initial investment of shale gas mining is relatively large, so the proper method for cutting down the investment is crucial. Among the various investments, the one related to the pipe network in shale gas gathering systems can account for more than 30% due to the long transmission distance (Arredondo-Ramirez et al., 2016; Zhou et al., 2019a). Therefore, this work aims to optimize the pipe network to make the system more economic and efficient.

Shale gas gathering system optimization is an NP-hard problem (Zhou et al., 2014a). The problem is NP-hard when any problem from NP is polynomial-time reducible to this problem. The optimization of shale gas gathering system can be generally divided into three sub-problems: well group division, site selection of gas collection station and central processing facilities, and pipeline optimization. For the well group division, Zhou et al., (2014b) established a mixed integer linear programming (MILP) model to optimize the well group division, the model is constrained to the number of wells connected to a gathering station, the allowable distance between well site and gathering station, and the number of gathering stations. Then, Zhou et al., (2019b) proposed an optimization scheme based on Delaunay triangulation to solve a real case study. Moreover, K-means algorithm is commonly used (Li et al., 2015; Qiang et al., 2016) to solve the problem of well clustering. Although it cannot constrain the number of gas-gathering station connections, it can simply and efficiently obtain the clustering results and is suitable for the large-scale classification problem. Its clustering process is sorted out by Larsson (2018). For

* Corresponding author.
E-mail address: wangyufei@cup.edu.cn (Y.-F. Wang).

https://doi.org/10.1016/j.petsci.2021.12.029

the site selection, Kerzmann et al., (2014) used Monte Carlo algorithm to optimize the location of gas stations in Pittsburgh according to the flow demand. Hosseini and MirHassani (2015) formulated a two-stage stochastic model considering the uncertainty of traffic flow on the road network and developed a heuristic algorithm based on core set and greedy algorithm to optimize the gas station location in Arizona. Based on network data of gas stations and geographic information system (GIS) data, Lin et al., (2020) proposed a hybrid algorithm by using genetic algorithm (GA), greedy algorithm and simulated annealing algorithm (SA). The algorithm was used to determine the site selection of hydrogen stations in Beijing. For pipeline optimization, pipeline topology, three-dimensional terrain, hydraulic characteristics (Hong et al., 2019) and reliability are often studied in current research attempts. However, there is a lack of synchronization among optimization solutions for these problems. To reduce the complexity of the model, most researchers tend to focus on solving one or two such problems. Until now, many metaheuristics algorithms have been applied, including ant colony optimization (ACO) (Baeza et al., 2017), GA (Sanaye and Mahmoudimehr, 2013), particle swarm optimization (PSO) (Zhang et al., 2017a), etc. Chan et al., (2007) found that compared with exhaustive search algorithms, probabilistic search heuristic algorithms are more suitable for solving pipeline network optimization problems. Zhang et al., (2017a) considered the gathering radius, economic flowrate and terrain obstacles and took the construction and operation costs as objective functions in the study of undersea oil pipeline optimization. They came up with a MILP model and used the improved PSO to optimize a real case. Hong et al., (2018) introduced an integrated model to optimize a 3-D gathering system, coupling SA and Dijkstra algorithm to solve the model. A recently published work (Meng et al., 2021b) compared the performances of ten recent and popular metaheuristics algorithms through several engineering examples of reliability-based design optimization problems. The work pointed out that artificial bee colony, Multi-verse optimizer, grey wolf optimizer, and water cycle algorithm are the most effective algorithms. However, the application of these algorithms in engineering is relatively lacking. In comparison, GA is more widely used.

Topology optimization plays an important role in pipeline network optimization. It not only determines the investment and energy loss of the pipeline network but also has a great impact on the reliability of the system. In the work of Hong et al. 2018, 2019, the optimal pipeline network has two levels and the topology types are both star networks. Su et al., (2019) considered the tree topology network and loop topology network to improve the reliability and supply efficiency of the natural gas pipe network. They adopted NSGA-II algorithm to optimize the pipe network system with multiple objectives. Zhang et al., (2017b) optimized the pipeline layout of stellated pipeline network (star topology), cascade dendritic pipeline network (tree topology) and insertion dendritic pipeline network (Steiner tree topology). A MILP model was presented to optimize the oil-gas collection pipeline system. Wu et al., (2020a) introduced the GeoSteiner algorithm to obtain the cable layout with topologies of Euclidean Steiner minimum tree (ESMT) and rectilinear Steiner minimum tree (RSMT) to solve the cable routing problem in wind farm layout. They emphasized that the cable length of these two structures is shorter than that of the minimum spanning tree (MST) topology. However, the influence of different pipeline network topologies on optimization has not been studied in all the mentioned researches. Therefore, this paper introduces two-level topology with MST, STAR, ESMT and RSMT simultaneously into the optimization and explores the influence of the combination of different topological structures on the objective function.

In addition, many pieces of research (Cen et al., 2016; Yu et al. 2018, 2019) incorporate economy and reliability into the network topology evaluation system. The economic performance (investment and operation cost) (Sanaye and Mahmoudimehr, 2013; Baeza et al., 2017) is usually considered as the objective function. There are also a few optimizations aimed at the environmental performance (Chen et al., 2020). To consider reliability, Takagi et al., (2002) studied the reliability of risk management of LNG receiving stations and pipe networks by using fault tree analysis. Recently, Alsharqawi et al., (2020) applied this method to the water distribution network. Liu and Pan (2003b) introduced flow reliability and unit importance in the reliability evaluation of tree-shaped subsea oil and gas gathering and transportation networks and calculated the reliability via Monte Carlo simulations. The authors in (Liu and Pan, 2003a) also optimized the system with a Minty algorithm to obtain the MST pipe network with the lowest cost and a certain reliability index. Based on graph theory, Christodoulou et al., (2018) used a Monte Carlo-based method to calculate the seismic reliability of urban water distribution network under the condition of given failure probability of pipeline and topology structure. Meanwhile, compared with the path enumeration method, the proposed method was more universal when the network was expanded. Kim and Kang (2013) developed a non-simulation network reliability analysis method based on a recursive decomposition algorithm. This algorithm is suitable for rapid risk assessment, but it needs additional implementation costs for defining complex system events and high computational costs for large networks with redundant paths. In the current research, there is a lack of study for the consideration of earthquakes in pipeline network optimization. Yoon et al., (2018) proposed a comprehensive framework to quantify the impact of earthquakes on the connectivity of urban water transmissions. The framework is a pipeline failure probability model based on the parameters of seismic depth, magnitude, and epicenter distance. The performance of the network topology was presented through Monte Carlo simulations. Unfortunately, their work aimed to evaluate the pipeline network considering earthquakes without further optimization. Thus, this research takes economic performance as the goal, and couples the reliability optimization of pipeline network to the gathering system optimization.

Since there are many uncertain factors in the production and transportation of shale gas, it is necessary to consider the uncertainty for the optimization problem. The uncertainties in optimization models could be described through bounded forms, probability description and fuzzy description (Li and Ierapetritou, 2008). Modeling methods mainly include robust programming, two-stage or multistage stochastic programming, and data-driven approaches that generate different possible scenarios (Grossmann et al., 2016). Meng et al., (2015) have done some meaningful work on reliability assessment methods considering uncertainty. For the evaluation of probabilistic constraints, they developed a modified chaos control method to improve the effectiveness of convex performance measure functions. In 2017, they (Meng et al., 2017) revealed the essential cause of low efficiency for the stability transform method and proposed a directional stability transformation method to control the non-convergence problems of the first-order reliability method. Recently, they (Meng et al., 2021a) presented a novel hybrid time-variant reliability model based on a probabilistic model and a super parametric convex model. The expansion optimal linear estimation method was used to discretize the stochastic process. The model was solved the model by a new relaxed reliability iteration method. For the uncertainty factors, Adachi and Ellingwood (2008) proposed a method considering the uncertainty of earthquake intensity for evaluating the serviceability of a water distribution system. The probabilities of peak ground accelerations and peak ground velocities are produced by a risk assessment platform (HAZUS). In their work, a region was divided into many sub-regions, which have different probabilities. In other words, the uncertainty is described by discrete parameters. Chen

et al., (2018) used life cycle analysis, interval linear programming, multi-objective programming, and multi-standard decision analysis to study the uncertainty of economy and environmental performance. The proposed method was used for the optimization of shale gas supply chain in the Marcellus region of the United States. Based on the beta distribution generated by historical data, Oke et al., (2020) studied the price uncertainty of shale gas using stochastic programming. Most published studies focused on the price uncertainty in the shale gas market, but the pipe failure uncertainty under no-earthquake condition and earthquake-related uncertainties (location, depth and magnitude) have been ignored. Therefore, the proposed model includes the uncertainties of the pipe failure and earthquake simultaneously.

To compensate for the deficiencies existing in present studies, this work proposes a method to optimize shale gas gathering system with the considerations of economic, pipeline network topology, reliability, earthquake-related uncertainty, pipe failure under no-earthquake conditions, and hydraulic characteristics. A chance-constrained coupling model is developed, and the objective function includes the system investment cost, operating cost and the loss of revenue caused by pipeline failure. GA is used to optimize the number of well clusters, site selection of gas collecting station and central processing facilities, pipe network structure, location, and the number of spare pipelines. In addition, Monte Carlo simulations, K-means algorithm, GeoSteiner algorithm, Kruskal algorithm and GLPK solver were used to solve different subproblems. In this paper, a real case is optimized to prove the effectiveness of the proposed method. The research also studies the influence of the number of clusters, spare pipelines and four kinds of pipeline topology on the shale gas system. The results are conducive to providing effective guidance for practical engineering design.

The main contributions of this paper are as follows:

(1) For the first time, the uncertainties of pipeline failure under no-earthquake conditions, earthquake location, magnitude and depth are considered simultaneously in the optimization of shale gas gathering system.
(2) The two-level pipeline network with four kinds of topologies is optimized in this research. The topologies include MST, STAR, ESMT and RSMT. Besides, the influence of different topology combinations on the whole system is also studied.
(3) This work not only optimizes network topology but also optimizes the location and quantity of spare pipelines to enhance the reliability of the gathering system.
(4) The proposed model considers the hydraulic characteristics of pipeline networks, pressure loss is taken to be the operation cost in the objective function.
(5) A variety of algorithms are integrated to optimize the shale gas gathering system comprehensively, and the proposed method can solve the practical large-scale optimization problem.

The study is outlined as follows: Section 1 reviewed the research progress related to the gathering system and explains the motivation for the proposed problem. The problem statement and mathematical model are described in detail in Section 2. Section 3 gives the flowchart of the proposed hybrid algorithm and introduces the application of different algorithms in the proposed optimization. Section 4 introduces the case study including scenario description and data preprocessing. Section 5 summarizes and discusses the results of scenarios with different models and sensitivity analyses. Finally, conclusions and future works are drawn in Section 6.

2. Methodology

2.1. Problem statement

The proposed method focuses on the optimization problem of the shale gas gathering system considering reliability enhancement under uncertainties. The problem consists of three sub-problems: well clustering, site selection for the gathering gas station (GGS) and central processing facility (CPF), and optimization of pipeline network topology considering no-earthquake uncertainty and earthquake-related uncertainties. The uncertainty under no-earthquake conditions denotes pipeline failures caused by improper operation, pipeline corrosion, and aging. Reliability is the performance in the overall pipeline gathering capacity when pipelines cannot function properly. In this research, the reliability under no-earthquake conditions is defined as conventional reliability, while the reliability under earthquake conditions is defined as earthquake reliability. Reliability is enhanced by increasing the number of spare pipelines and optimizing their locations.

To describe the proposed optimization problem, a coupling model is developed, as shown in Fig. 1. The outer model is used to optimize the layout and reliability of the gathering system. Specifically, the decision variables of the model are cluster number K, tree types of different levels (first level and second level) of pipeline network, number, cluster and position (first node and second node) of spare pipelines, positions (nodes) of CPF and GGS. The input data are coordinates and gas production of wells, earthquake data and other related parameters of formulas of the model. The objective function is to minimize the total annual cost *TAC* (CNY·a^{-1}). The inner model is used for clustering and network topology optimization. There is a linear programming (LP) model to calculate the fluid flow in the network. Uncertainties are described by a chance-controlled model. The details of the proposed models are shown in Section 2.2.

2.1.1. Assumptions

(1) The shale gas is regarded as a compressible fluid.
(2) The shale gas in each branch has the same inlet pressure.
(3) The temperature of the shale gas in each branch is constant.
(4) The flow rate of the shale gas in each branch is constant.
(5) The probability of conventional failure under no-earthquake conditions per unit length of the pipeline is fixed.

2.1.2. Given

(1) The coordinates of each shale gas well.
(2) The shale gas production of each well.
(3) Historical seismic data in the studied area.

2.1.3. Determine

(1) The number of well clusters (K).
(2) The location of the GGS and CPF.
(3) The number and location of spare pipelines.
(4) The topology of the collection and transmission network.
(5) The flow rate in each pipeline and the diameter of each pipeline.
(6) The total annual cost of the shale gas gathering system.
(7) Conventional reliability and earthquake reliability of the pipeline network.

According to the problem statement, it can be summarized the

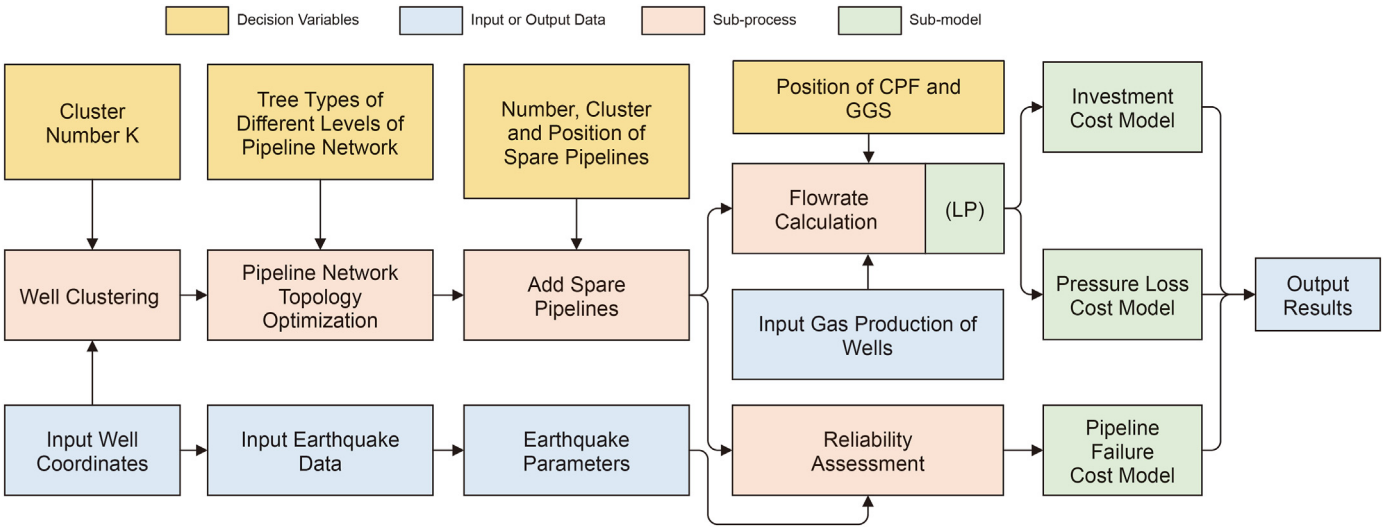

Fig. 1. Coupling model for the optimization of shale gas gathering system.

main contributions of the proposed methodology as follows:

(1) The proposed methodology can help achieve the synchronization optimization of well clustering, site selection for GGS & CPF, and optimization of pipeline network topology.
(2) The proposed methodology considers conventional reliability and earthquake reliability at the same time.
(3) The optimization of the spare pipeline (number and position) and tree type can be used to enhance the reliability of the pipeline network. The reliability assessment for the network with loops is solved by the proposed methodology.
(4) The hydraulic characteristics of pipeline networks, pressure loss, is taken to consideration.
(5) The proposed methodology can describe the optimization of a hybrid pipeline network with different levels.

2.2. Mathematical Model

2.2.1. Objective function

The objective function is minimizing *TAC*. It consists of the construction cost of CPF C^{CPF} (CNY·a^{-1}), the construction cost of GGS C_k^{GGS} (CNY·a^{-1}), the construction cost of the pipeline $C_{(i,j)}^{\mathrm{Pipeline}}$ (CNY·a^{-1}), the pressure loss cost of the pipeline $C_{(i,j)}^{\mathrm{Pressure}}$ (CNY·a^{-1}), and the expectation of the loss cost of pipeline failures $C^{\mathrm{Reliability}}$ (CNY·a^{-1}), as shown in Eq. (1).

$$\min TAC = C^{\mathrm{CPF}} + \sum_k C_k^{\mathrm{GGS}} + \sum_{(i,j)} C_{(i,j)}^{\mathrm{Pipeline}} + \sum_{(i,j)} C_{(i,j)}^{\mathrm{Pressure}} + C^{\mathrm{Reliability}}, \quad \forall (i,j) \in \mathbf{E_T},\ k \in \mathbf{V^{GGS}} \tag{1}$$

2.2.1. Constraints.

$$1 \le n^{\mathrm{K}} \le \mathrm{n}^{\mathrm{K,max}} \tag{2}$$

$$0 \le T_{\mathrm{first}}^{\mathrm{tree}} \le \mathrm{T}_{\mathrm{max}}^{\mathrm{tree}} \tag{3}$$

$$0 \le T_{\mathrm{second}}^{\mathrm{tree}} \le \mathrm{T}_{\mathrm{max}}^{\mathrm{tree}} \tag{4}$$

$$0 \le n^{\mathrm{ASP}} \le \mathrm{n}^{\mathrm{ASP,max}} \tag{5}$$

Eqs. (2)–(5) are the upper and lower bound constraints of the variables, n^{K} is the number of well clusters clustered, $T_{\mathrm{first}}^{\mathrm{tree}}$ is the topology type of the first level of pipeline network, $T_{\mathrm{second}}^{\mathrm{tree}}$ is the topology type of the second level of pipeline network, and n^{ASP} is the number of spare pipelines. These variables are integer variables.

2.2.2. Linear programming model

The mass flow rate of shale gas in each branch pipe of each level of the pipe network is obtained by solving an LP model (Wu et al., 2020b). The model is established based on the results of the pipe topology optimization. The model is described by Eqs. (6)–(9).

$$\min W = \sum_{\gamma}^{n^{V_T}} \sum_{\alpha}^{n^{V_T}} \sum_{\beta}^{n^{V_T}} b_{\alpha,\beta,\gamma}^{\mathrm{Binary}} \times w_{\alpha,\beta,\gamma} \tag{6}$$

$$\sum_{\alpha}^{n^{V_T}} \sum_{\beta}^{n^{V_T}} \left[(-1)^D b_{\alpha,\beta,\gamma}^{\mathrm{Binary}} \times w_{\alpha,\beta,\gamma} + (-1)^D b_{\alpha,\beta,\gamma}^{\mathrm{Binary}} \times w_{\beta,\alpha,\gamma} \right] - W_\gamma = 0 \tag{7}$$

$$w_{\alpha,\beta,\gamma} \ge 0 \tag{8}$$

$$w_{\beta,\alpha,\gamma} \ge 0 \tag{9}$$

The constraints are shown as Eqs. (7)–(9). Eq. (7) is used to calculate the material balance for each vertex, Eqs. (8)–(10) are used to control the gas flow direction. The decision variable is $w_{\alpha,\beta,\gamma}$ (kg·s^{-1}), denoting the mass flow rate of shale gas in the edge (α, β) connected to the vertex γ when the material balance is performed on the vertex γ. The remaining symbols represent parameters of the LP model. n^{V_T} is the number of vertices in the pipeline network topology, W_γ (kg·s^{-1}) is the mass flow rate of the vertex γ, if vertex γ produces gas, W_γ is negative, otherwise, W_γ is positive. The flow rate in each pipeline has two directions, from α to β and from β to α. In Eq. (7), for each vertex γ, there is a set of edges $\mathbf{E_T}$ connected to it, each set $\mathbf{E_T}$ is a subset of the connected tree, and the number of edge sets is n^{V_T} so that one of the two vertices corresponding to

each edge $\mathbf{E_T}$ must be γ. If the edge (α, β) belongs to $\mathbf{E_T}$, $b^{\text{Binary}}_{\alpha,\beta,\gamma}$ is equal to 1, otherwise, it is equal to 0. The mass flow rate of edge (α, β) is calculated twice from α to β ($w_{\alpha,\beta,\gamma}$) and β to α ($w_{\beta,\alpha,\gamma}$). These two mass flow rates are independent of each other due to the constraints (Eq. (8) and Eq. (9)). After solving the optimization, one of $w_{\alpha,\beta,\gamma}$ and $w_{\beta,\alpha,\gamma}$ must be 0, but they will not be 0 at the same time, and the one that is not 0 is a positive value. Besides, $b^{\text{Binary}}_{\alpha,\beta,\gamma}$ is the binary parameter of the model and obtained from Eq. (10) based on the result of the optimized pipeline network topology, $b^{\text{Binary}}_{\alpha,\beta,\gamma}$ indicates whether there is an edge (α, β), which belongs to the connected tree, directly connected to the vertex γ. D is the exponential parameter and is determined by Eq. (11).

$$b^{\text{Binary}}_{\alpha,\beta,\gamma} = \begin{cases} 0, & [\alpha \neq i] \vee [\beta \neq j] \\ 1, & [\alpha = i] \wedge [\beta = j] \end{cases}, \quad \forall (i,j) \in \mathbf{E_T} \tag{10}$$

$$D = \begin{cases} 1, & \alpha < \beta \\ 2, & \alpha \geq \beta \end{cases} \tag{11}$$

2.2.3. Investment cost model

The construction cost of the CPF C^{CPF} (CNY·a^{-1}), the construction cost of the GGS C^{GGS}_k (CNY·a^{-1}), the construction cost of the pipeline $C^{\text{Pipeline}}_{(i,j)}$ (CNY·a^{-1}), and the unit cost of the pipeline $a^{\text{Pipeline}}_{(i,j)}$ (CNY·a^{-1}) (Stijepovic and Linke, 2011) are calculated from Eqs. (12)–(15), respectively.

$$C^{\text{CPF}} = \frac{(1+I)^T I}{(1+I)^T - 1} a^{\text{CPF}} \tag{12}$$

$$C^{\text{GGS}}_k = \frac{(1+I)^T I}{(1+I)^T - 1} a^{\text{GGS}}_k, \quad \forall k \in \mathrm{V}^{\mathbf{GGS}} \tag{13}$$

$$C^{\text{Pipeline}}_{(i,j)} = \frac{(1+I)^T I}{(1+I)^T - 1} a^{\text{Pipeline}}_{(i,j)} L_{(i,j)}, \quad \forall (i,j) \in \mathbf{E_T} \tag{14}$$

$$a^{\text{Pipeline}}_{(i,j)} = 5.74 Wt_{(i,j)} + 1295 \left(D^{\text{out}}_{(i,j)}\right)^{0.48} + 47.6, \quad \forall (i,j) \in \mathbf{E_T} \tag{15}$$

$$Wt_{(i,j)} = 644.3 \left(D^{\text{in}}_{(i,j)}\right)^2 + 72.5 D^{\text{in}}_{(i,j)} + 0.4611, \quad \forall (i,j) \in \mathbf{E_T} \tag{16}$$

$$D^{\text{out}}_{(i,j)} = 1.052 D^{\text{in}}_{(i,j)} + 0.005251, \quad \forall (i,j) \in \mathbf{E_T} \tag{17}$$

$$D^{\text{in}}_{(i,j)} = \sqrt{\frac{4W_{(i,j)}}{\pi \rho u_{(i,j)}}}, \quad \forall (i,j) \in \mathrm{E}_{\mathbf{T}} \tag{18}$$

where T (a) is the life cycle of the piping system, I is the annual interest rate, and $L_{(i,j)}$ (m) is the length of the pipeline. The weight per unit length of the pipeline $Wt_{(i,j)}$ (kg·m^{-1}), pipeline outer diameter $D^{\text{out}}_{(i,j)}$ (m) and pipeline inner diameter $D^{\text{in}}_{(i,j)}$ (m) are calculated with Eqs. (16)–(18), respectively, where $W_{(i,j)}$ (kg·s^{-1}) is the mass flow rate of the fluid in the pipeline, ρ (kg·m^{-3}) is the density of the fluid in the pipeline, and $u_{(i,j)}$ (m·s^{-1}) is the flow rate of the fluid in the pipeline.

2.2.4. Pressure loss cost model

$$C^{\text{Pressure}}_{(i,j)} = \frac{a^{\text{Electricity}} t N_{(i,j)}}{1000}, \quad \forall (i,j) \in \mathbf{E_T} \tag{19}$$

$$N_{(i,j)} = \frac{Ne_{(i,j)}}{\eta}, \quad \forall (i,j) \in \mathbf{E_T} \tag{20}$$

$$Ne_{(i,j)} = H^{\text{f}}_{(i,j)} W_{(i,j)} g, \quad \forall (i,j) \in \mathbf{E_T} \tag{21}$$

The pressure loss cost $C^{\text{Pressure}}_{(i,j)}$ (CNY·a^{-1}) of the pipeline network, the shaft power $N_{(i,j)}$ (W) to overcome the resistance when the gas is transported and the effective power $Ne_{(i,j)}$ (W) of the conveying equipment are calculated by Eqs. (19)–(21), respectively, where $a^{\text{Electricity}}$ (CNY·kW^{-1}·h^{-1}) is the electricity cost, t (h) is the annual operating hours, η is the efficiency of the conveying equipment, and $H^{\text{f}}_{(i,j)}$ (m) is the head loss due to the resistance along the route.

$$\frac{\rho \left[\left(P^{\text{in}}_{(i,j)}\right)^2 - \left(P^{\text{out}}_{(i,j)}\right)^2 \right]}{2P^{\text{in}}_{(i,j)}} \geq G^2_{(i,j)} \left\{ \ln \frac{P^{\text{in}}_{(i,j)}}{P^{\text{out}}_{(i,j)}} + \frac{\zeta}{2} \frac{L_{(i,j)}}{D^{\text{in}}_{(i,j)}} \right\} \times 10^{-3}, \quad \forall (i,j) \in \mathbf{E_T} \tag{22}$$

$$H^{\text{f}}_{(i,j)} = \frac{P^{\text{in}}_{(i,j)} - P^{\text{out}}_{(i,j)}}{\rho g}, \quad \forall (i,j) \in \mathbf{E_T} \tag{23}$$

$H^{\text{f}}_{(i,j)}$ (m) can be calculated iteratively according to Eqs. (22) and (23) (Wu et al., 2020b) regarding the standard guideline for pipeline sizing in the petrochemical industry (SH/T3035-2018). ρ (kg·m^{-3}) is the density of the fluid in the pipeline, $P^{\text{in}}_{(i,j)}$ (kPa) is the inlet pressure of the pipeline, $P^{\text{out}}_{(i,j)}$ (kPa) is the outlet pressure of the pipeline, ζ is the coefficient of friction of the pipeline.

2.2.5. Reliability evaluation and pipeline failure cost model

$$C^{\text{Reliability}} = \left\{ \begin{array}{l} \left[1 - E\left(R^{\text{Conventional}}\right)\right]\left(1 - P^{\text{Earthquake}}\right) \\ + \left[1 - \overline{R}^{\text{Earthquake}}\right] P^{\text{Earthquake}} \end{array} \right\} a^{\text{Shale gas}} \sum_i Q_i, \quad \forall i \in \mathbf{V_T} \tag{24}$$

$C^{\text{reliability}}$ indicates the pipeline failure cost and is related to the reliability of the pipeline network. It is calculated by Eq. (24) and includes two parts, one is the pipeline failure cost under no-earthquake conditions (conventional failure cost) and the other is the pipeline failure cost in the event of the earthquake (earthquake failure cost). $a^{\text{Shale gas}}$ (CNY·m^{-3}) is the unit price of shale gas, $P^{\text{Earthquake}}$ is the probability of an earthquake, Q_i (kg·s^{-1}) is the production of shale gas per well, $E(R^{\text{Conventional}})$ and $\overline{R}^{\text{Earthquake}}$ are the reliabilities of the pipeline network system under no-earthquake conditions and in the event of the earthquake, respectively, where pipeline failure under conventional conditions may be due to pipeline corrosion, aging, or improper operation.

$$E\left(R^{\text{Conventional}}\right)=\frac{\sum_{l}\sum_{i}q_{l,i}}{n^{S}\sum_{i}Q_i},\quad \forall i\in\mathbf{V_T},l\in\mathbf{N^L} \tag{25}$$

$$q_{l,i}\left(or\ q_{l,s,i}\right)=\begin{cases}0, & m_{i,\text{CPF}}=0\\ Q_i, & m_{i,\text{CPF}}\geq 1\end{cases},\quad \forall i,\text{CPF}\in\mathbf{V_T} \tag{28}$$

$$\mathrm{M}_{n^{V_T}\times n^{V_T}}=\bigcup_{k=1}^{n^{V_T}}\mathrm{A}^{(k)}_{n^{V_T}\times n^{V_T}} \tag{29}$$

$$\mathrm{A}^{(k)}_{n^{V_T}\times n^{V_T}}=\mathrm{A}^{(k-1)}_{n^{V_T}\times n^{V_T}}\odot\mathrm{A}_{n^{V_T}\times n^{V_T}} \tag{30}$$

$m_{i,CPF}$ is an element of the reachability matrix $\mathrm{M}_{n^{V_T}\times n^{V_T}}$, which is equal to 1 when there is at least one path between the vertex i and the vertex which indicates CPF, and 0 otherwise. $\mathrm{M}_{n^{V_T}\times n^{V_T}}$ can be obtained by performing Boolean addition and Boolean multiplication on the adjacency matrix $\mathrm{A}^{(k)}_{n^{V_T}\times n^{V_T}}$ by Eqs. (29)-(30) (Wu et al., 2021). In Eq. (30), $\odot$ is a mathematical operation symbol used for logical operations, representing "or".

$$a_{i,j}^{\text{Conventional}}=\begin{cases}1, & \left[r_{i,j}^{\text{Conventional}}>\left(1-\left(RR^{\text{Conventional}}\right)^{\frac{L_{i,j}}{1000}}\right)\right]\cap[(i,j)\in\mathrm{E_T}]\\ 0, & \text{otherwise}\end{cases},\quad \forall i,j\in\mathbf{V_T} \tag{31}$$

$$\left(a_{i,j}^{\text{Earthquake}}\right)_{(M,H,R)}=\begin{cases}1, & \left[r_{i,j}^{\text{Earthquake}}>\left(1-\left(RR^{\text{Conventional}}\right)^{\frac{L_{i,j}}{1000}}\times e^{\frac{-RR_{i,j}^{\text{Earthquake}}L_{i,j}}{1000}}\right)\right]\cap[(i,j)\in\mathbf{E_T}]\\ 0, & \text{otherwise}\end{cases},\quad \forall i,j\in\mathbf{V_T} \tag{32}$$

$$\begin{aligned}&\min\overline{R}^{\text{Earthquake}}\\ &s.t.\\ &\Pr\left\{R_l^{\text{Earthquake}}(M,H,R)\geq\overline{R}^{\text{Earthquake}}\right\}\geq c^{\text{Earthquake}}\end{aligned} \tag{26}$$

$$L_{i,j}=\begin{cases}L_{(i,j)}, & \forall(i,j)\in\mathbf{E_T}\\ 0, & \text{otherwise}\end{cases} \tag{33}$$

$$R_l^{\text{Earthquake}}(M,H,R)=\frac{\sum_{s}\sum_{i}\left(q_{l,s,i}\right)_{(M,H,R)}}{n^{S}\sum_{i}Q_i},\quad \forall i\in\mathbf{V_T},s\in\mathbf{N^S},l\in\mathbf{N^L} \tag{27}$$

$$RR_{i,j}^{\text{Earthquake}}=\begin{cases}RR_{(i,j)}^{\text{Earthquake}}=0.050\left(v_{(i,j)}^{\text{scaled}}\right)^{0.865}, & \forall(i,j)\in\mathbf{E_T}\\ 0, & \text{otherwise}\end{cases} \tag{34}$$

$$v_{(i,j)}^{\text{scaled}}=\frac{10^5\times PGV_{(i,j)}}{\left(D_{(i,j)}^{\text{out}}\times 100\right)^{1.138}},\quad \forall(i,j)\in\mathbf{E_T} \tag{35}$$

$$\log_{10}\left(PGV_{(i,j)}\right)=\begin{cases}0.725M+0.00318H-0.519\\ -1.318\times\log_{10}\left(R_{(i,j)}+0.334e^{0.653M}\right)\end{cases},\forall(i,j)\in\mathbf{E_T} \tag{36}$$

$E(R^{\text{Conventional}})$ and $\overline{R}^{\text{Earthquake}}$ are calculated by Eqs. (25)–(27), respectively, where n^S is the number of Monte Carlo simulations for different parameters of the earthquake, $q_{l,i}$ is the shale gas flowrate from ith vertex to the CPF in the lth simulation for the location of pipe failure under the conventional conditions. $q_{l,s,i}$ is the shale gas flowrate from ith vertex to the CPF in the lth simulation for the location of the earthquake and sth simulation for the seismic depth and magnitude under earthquake conditions. Eq. (26) describes that, in Monte Carlo simulations, the probability of all $R_l^{\text{Earthquake}}$ greater than $\overline{R}^{\text{Earthquake}}$ is greater than $c^{\text{Earthquake}}$. $c^{\text{Earthquake}}$ is the confidence level of reliability under earthquake conditions.

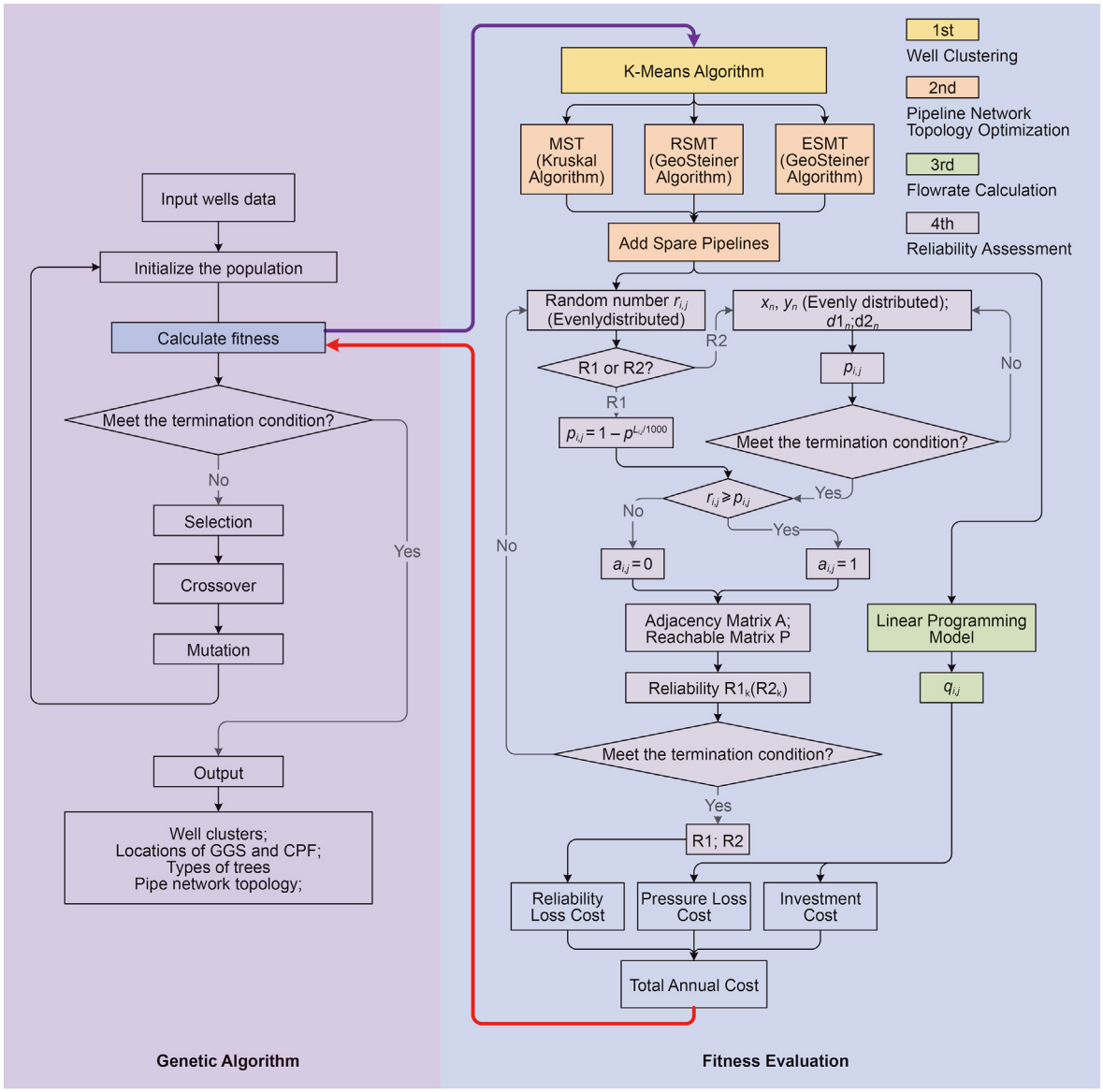

Fig. 2. Flow diagram of algorithms.

$$p_{(i,j)}^{\text{Depth}}(H) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(H-\mu)^2}{2\sigma^2}\right) \tag{37}$$

$$p_{(i,j)}^{\text{Magnitude}}(M) = a \times \exp(-b \times M), \quad \forall (i,j) \in \mathbf{E_T} \tag{38}$$

Eqs. (31)–(32) are used to determine if the uncertainty factors (pipeline failure and earthquake) are realized by comparing the random number generated by Monte Carlo with the probability of occurrence. Besides, the length of the pipeline in these two equations is obtained by Eq (33). Eqs. (34)–(36) are used to evaluate the damage to the pipeline caused by the earthquake. The probability distribution of the depth and magnitude of the earthquake in the Monte Carlo simulation is expressed by Eqs. (37)–(38).

In Eqs. (31)-(36), $a_{i,j}^{\text{Conventional}}$ and $a_{i,j}^{\text{Earthquake}}$ are the elements of $A_{n^v \times n^v}$, $RR^{\text{Conventional}}$ (km^{-1}) is the unit reliability of the pipeline under conventional conditions, $RR_{i,j}^{\text{Earthquake}}$ (Tromans, 2004) is the repairing rate of the branch pipeline (i, j) in the event of an earthquake, $v_{(i,j)}^{\text{scaled}}$ (Tromans, 2004) is the proportional velocity on the edge (i, j), $PGV_{(i,j)}$ ($km \cdot s^{-1}$) (Yoon et al., 2018) is the peak ground velocity, $D_{(i,j)}^{\text{out}}$ is the outer diameter of the pipeline, M is the seismic magnitude, H (km)is the depth of the source, and $R_{(i,j)}$ (km) is the distance from the pipeline (i, j) to the epicenter.

In Eqs. (37)-(38), $p_{(i,j)}^{\text{Depth}}$ is the probability of seismic depth H when the earthquake occurs, and $p_{(i,j)}^{\text{Magnitude}}$ is the probability of seismic magnitude M when the earthquake occurs. The two probability functions are fitted from historical seismic data.

3. Algorithms

To solve the proposed model, this study combines GA, a clustering algorithm, a graph theory-based algorithm, and Monte Carlo simulations to optimize the layout of the shale gas surface gathering system. The framework of the algorithm is shown in Fig. 2. The algorithm includes two parts: the inner layer and the outer

layer. The outer layer uses GA to determine the number of well clusters, the location variables of the CPF and GGS, the number and location of spare pipelines, and topology type (STAR, MST, ESMT, and RSMT) for the two-level pipeline network. The inner layer focus on the calculation of the fitness function. It consists of four parts: (1) The K-means algorithm is used for well group division. (2) Depending on topology types of different levels of pipeline network, Kruskal or GeoSteiner algorithm is applied to optimize the topology for the GGS (first level pipeline network) and shale gas wells (second level pipeline network). Then the spare pipelines are added to the optimized loop-free pipe network. (3) Solve the LP model to obtain the flow rate of shale gas in each pipeline and perform hydraulic calculations. (4) Calculate the conventional reliability and earthquake reliability of pipe network with loops using reachability matrix and Monte Carlo simulations.

3.1. Genetic algorithm

GA is a global probabilistic algorithm based on the process of genetics and natural selection (Roetzel et al., 2020). It can find optimal or near-optimal solutions to constrained and unconstrained problems. The algorithm simulates the evolutionary process of organisms in nature and follows the rules of natural selection and survival of the fittest. GA has good robustness in solving complex optimization problems and is widely used in solving network system problems. Besides, GA is usually coupled with the mathematical model in published researches because of its good versatility. Cariou et al., (2018) solved a MILP model by CPLEX and a specific GA to solve the liner shipping network design problem. It was illustrated that their GA can reach an optimal solution when solving large-size instances. Wu and Wang (2021) developed a generally applicable framework to optimize the recovery strategies of the network system. Mathematical programming was used to achieve the coordination of different recovery agents and GA was used to find the assignment of the required resource for restoration. In the work of Nemati et al., (2018), an improved real-coded GA and enhanced MILP based method were proposed to solve the dispatching optimization problem of microgrids. Compared with other new algorithms that are more efficient in solving individual cases, the effectiveness of GA has been proven in the application of many engineering cases. Choosing GA is a safer method that can get good results. Therefore, GA is selected used for solving the gathering system optimization considering piping network and facility layout in this research.

The GA process in this study is shown in Fig. 2: firstly, the population is randomly initialized; secondly, the finesses functions of all the individuals are calculated to evaluate the quality of the solution; thirdly, if the termination condition is not met, a new generation of the population is generated by selection, crossover and mutation of elite individuals in the parent generation; finally, the best individual survives the continuous evolutionary process, and the optimized solution can be obtained after decoding.

Fig. 3 displays the physical meanings of all decision variables of the mathematical model. The variables are represented by genes in GA. Specifically, the variables include the number of well clusters, the topology type of the pipeline network, the number and location of spare pipelines, and the locations of GGSs and CPF.

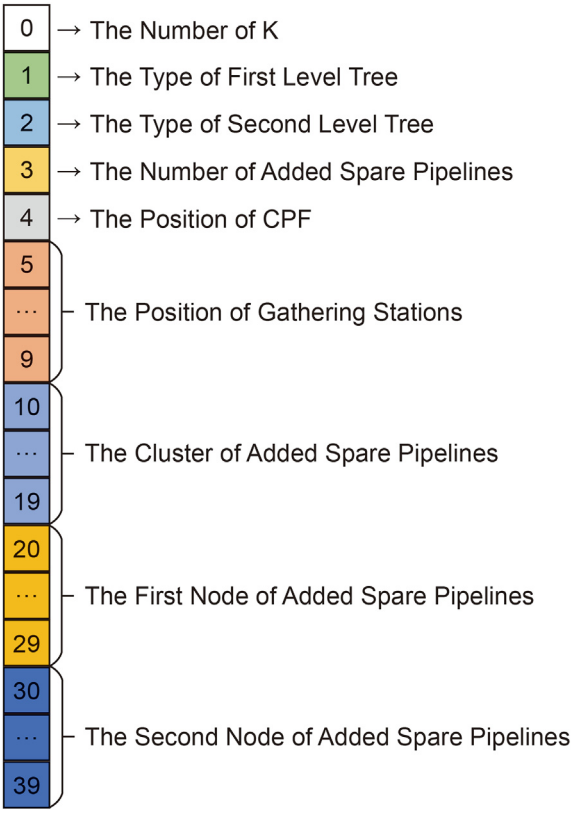

Fig. 3. The chromosome of GA.

3.2. K-means algorithm

The K-means algorithm is a clustering algorithm. It can classify data objects with high similarity into the same class clusters and classify data objects with high dissimilarity into different class clusters. The specific process of the K-means algorithm in this paper is as follows.

(1) Select K points as the initial centroids.
(2) Assign each point to the nearest centroid to form K clusters and recalculate the centroid of each cluster.
(3) Repeat steps (2) and (3) until there is no centroids change.

3.3. Kruskal algorithm

Kruskal algorithm and Prim algorithm are classical algorithms for solving the MST of a weighted connected graph (Li et al., 2017). Kruskal algorithm is best for very sparse graphs whereas the Prim algorithm is fastest for denser graphs but requires more memory (Dementiev et al., 2004). The topology in this research is optimized in a sparse graph, so Kruskal algorithm is selected to solve the MST in this research. It finds the edge with the smallest weight, and then gradually turns the connected component into an MST that connects all vertices. The specific process of Kruskal algorithm is as follows.

(1) Given a weighted connected graph **G (V, E, W)** of a pipeline network. **V**, **E** and **W** are the set of vertices, edges and weights in the connected graph, respectively. Initially, the set of edges in MST is empty.
(1) Sort edges in **E** from smallest to largest according to the value of $w_{i,j}$ in **W** to obtain e_1, e_2, e_3, … e_{N_q}.
(2) Add edges to MST ($\mathbf{E_T}$) sequentially in sequence and make a judgment. If the addition does not form a loop, keep the addition of e_1, otherwise discard e_1.
(3) If the number of edges in the MST is $N_p - 1$, stop the algorithm and output the result **T** ($\mathbf{V_T}$, $\mathbf{E_T}$, $\mathbf{W_T}$), otherwise, go to step (3). N_p is the number of vertices in the connected graph.

3.4. GeoSteiner algorithm

The GeoSteiner algorithm is currently a very efficient method for accurately solving different types of Steiner trees. It was improved by the work of Zachariasen (1998) and Warme (1998). Based on the algorithm, their team also developed a GeoSteiner software package. The program has for about 20 years been the fastest tool for computing exact solutions to Steiner tree problems in the plane(-Juhl et al., 2018). Steiner tree is a topology that connects given points in a plane and allows other points besides the given points to be included in the topology. The generation of Steiner tree consists of two very important steps: the generation and connection of full Steiner trees (FSTs). Before generating the FSTs, a relatively time-consuming preprocessing process is performed to reduce the complexity and to increase the speed of the calculation. The connection of FSTs is the bottleneck of GeoSteiner algorithm. This step can be treated as a problem of finding the MST in a hypergraph. It can be solved by integer programming or dynamic programming. In this paper, the GeoSteiner package is embedded in the optimization program and the integer programming is used to solve the problem.

3.5. Reliability assessment (Monte Carlo simulations)

Monte Carlo method is a numerical computational method guided by probability and statistics. It provides approximate solutions to engineering problems through statistical sampling experiments. The implementation of the method consists of two steps: (1) Use a large number of random variables that follow some probability distribution to simulate the real problem. (2) Use statistical methods to estimate the numerical characteristics of the model and to obtain the numerical solution of the problem. In this work, the pipelines network has loops due to the addition of spare pipelines, leading to a much more difficult calculation of reliability. The widely used matrix operations cannot accurately calculate the reliability of such a complex pipeline network, so Monte Carlo simulations are used in this research. Monte Carlo simulations are used to calculate conventional reliability and earthquake reliability respectively. When the network is in conventional operation, the pipeline failures are uniformly distributed random numbers. In the event of an earthquake, the failure of the pipeline is related to the location, depth, and magnitude of the earthquake. Besides, the numerical characteristic of the model is described by Eq. (26). All the solutions obtained from Monte Carlo simulation will be sorted from largest to smallest, then $\overline{R}^{\text{Earthquake}}$ will be the $c^{\text{Earthquake}} \times n^{S}$ th value in sorted solutions.

In the proposed model, each pipeline branch corresponds to an intermediate variable. The intermediate variable represents the total failure probability of this branch pipeline. The total failure probability is related to the unit failure probability of each pipeline, the pipeline length, the pipeline coordinates, and various parameters of the earthquake. As shown in Fig. 2, outer Monte Carlo simulations generate all the random numbers used to determine whether a pipeline has failed. In each simulation, when the pipeline network is operating normally (R1), the total probability of pipeline failure is obtained through the unit probability of failure and the pipeline length. When an earthquake occurs (R2), inner Monte Carlo simulations are required to obtain the earthquake-related random numbers. The random numbers follow their respective distributions and include earthquake location (uniform distribution), depth (normal distribution), and magnitude (exponential distribution). After the number of simulations is satisfied, the approximate solution of the total failure probability of the branch pipeline under earthquake conditions can be determined. After that, the random numbers and the total failure probability generated uniformly by the outer Monte Carlo simulations are used to determine whether the branch pipeline fails or not. In this way, an updated pipeline network without failed pipeline can be obtained. Based on these results, the final reliability can be obtained by calculating the ratio of the total amount of shale gas delivered by the updated pipeline network to that of the original network. When the number of simulations is large enough, the approximate solution of the conventional reliability is the expected value of reliabilities obtained by all simulations, and the earthquake reliability is the minimum of the reliabilities obtained by all simulations that meet the reliability constraints.

Table 1
The conditions of scenarios 1 to 4.

Scenario index	K-value	Pipeline topology	Added spare pipeline
Scenario 1	2	First level: MST Second level: MST	0
Scenario 2	Variable	First level: MST Second level: MST	0
Scenario 3	Variable	First level: Variable Second level: Variable	0
Scenario 4	Variable	First level: Variable Second level: Variable	Variable

4. Case study

4.1. Scenario description

A case study from a real shale gas project is used to illustrate the effectiveness of the proposed method. In this case study, 42 shale gas wells cover an area of about 8.5 hm^2, and the total annual output of shale gas reaches 1.851×10^6 m^3. This section includes four scenarios and three sensitivity analyses. The scenarios consist of an original scenario and three scenarios with different variables which are shown in Table 1. To further explore the impact of some important variables on the objective function, the sensitivity analyses on the impact of the numbers of clusters, topology type of pipeline network and spare pipelines number are carried out. The optimization is implemented in C++ by QT in Ubuntu Linux 14.04.2. There are 2 CPUs where each of which is Intel(R) Xeon(R) Gold 6226 R CPU @ 2.90 GHz. The Available random-access memory (RAM) is 128G.

4.2. Data acquisition and preprocessing

The initial data of the case study comes from a shale gas production area in China. The well data are shown in Table 2 including index, name, coordinates, and output of wells.

The earthquake data comes from the China Earthquake Network Center[1] (CENC). The data from 2010 to 2019 of the area where shale gas wells are located are collected. Figs. 4 and 5 display the fitting results of seismic depth and magnitude, the corresponding formulas are shown in Eqs. (39)-(40).

$$p_{(i,j)}^{\text{Depth}}(H) = 0.07792 \times \exp\left[-\left(\frac{H - 13.75}{7.468}\right)^2 \right], \quad \forall (i,j) \in \mathbf{E_T} \tag{39}$$

[1] http://www.ceic.ac.cn.

Table 2
Input data of wells.

Well Index	Well Name	Coordinate X, m	Coordinate Y, m	Well Output, 10^4 m^3/d
0	Well-0	9572.40	10748.90	3.2
1	Well-1	9993.90	9313.80	4.0
2	Well-2	10445.20	11305.00	5.0
3	Well-3	10649.40	12520.20	5.0
4	Well-4	1071 2.10	14296.20	5.0
5	Well-5	9651.80	7251.40	5.0
6	Well-6	11357.90	15601.40	5.0
7	Well-7	11401.00	17859.80	5.0
8	Well-8	12625.10	9901.00	5.0
9	Well-9	11656.10	12025.00	2.5
10	Well-10	12104.00	13668.90	3.2
11	Well-11	12577.10	15490.20	3.2
12	Well-12	12959.20	19734.00	5.0
13	Well-13	13435.30	9254.40	2.5
14	Well-14	13689.10	9433.00	5.0
15	Well-15	16115.10	10850.00	2.5
16	Well-16	15636.90	12177.60	5.0
17	Well-17	15527.10	13040.00	5.0
18	Well-18	16297.10	14207.00	5.0
19	Well-19	15785.10	15318.00	5.0
20	Well-20	15806.10	16635.00	5.0
21	Well-21	15713.10	17789.00	5.0
22	Well-22	16091.10	19050.00	5.0
23	Well-23	15699.10	21122.00	5.0
24	Well-24	6270.10	16258.00	2.5
25	Well-25	13058.10	20154.00	2.5
26	Well-26	6609.10	8897.00	6.0
27	Well-27	6553.20	9903.85	6.0
28	Well-28	7533.90	8058.80	3.0
29	Well-29	3879.74	7133.97	5.0
30	Well-30	6219.76	5366.60	2.5
31	Well-31	6862.10	3241.00	5.0
32	Well-32	5977.10	1283.00	5.0
33	Well-33	6722.10	0.00	5.0
34	Well-34	6667.10	1283.00	5.0
35	Well-35	6604.30	4318.70	3.0
36	Well-36	8259.19	6216.39	3.0
37	Well-37	1136.10	11102.00	5.0
38	Well-38	2514.10	12745.00	5.0
39	Well-39	1932.10	9629.80	5.0
40	Well-40	2332.88	8052.59	5.0
41	Well-41	0.00	7614.80	5.0

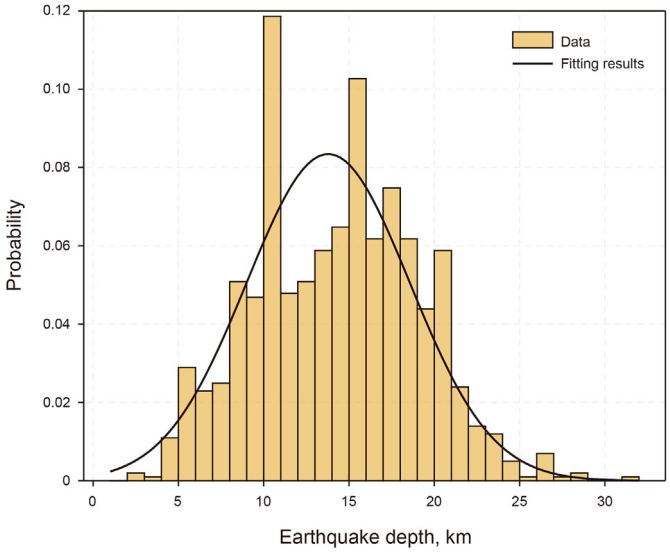

Fig. 4. Distribution of seismic depth.

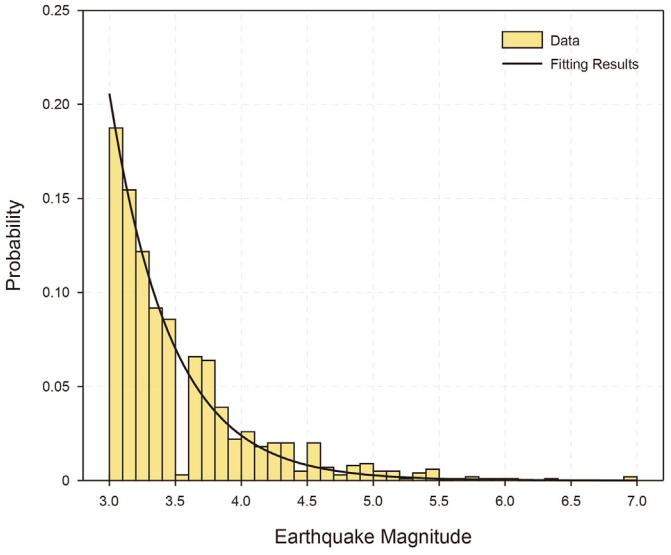

Fig. 5. Distribution of seismic magnitude.

Table 3
Input data of main parameters.

Genetic algorithm		Pressure loss model		Pipeline cost model		Reliability model	
Crossover	0.3	$a^{\text{Electricity}}$	0.21	C^{CPF}	2×10^7	$a^{\text{Shale gas}}$	2.22
Generation	400	t	8400	C^{GGS}	5.4×10^7	n^{S}	1000
Mutation	0.8	u	150	I	0.02	$c^{\text{Earthquake}}$	0.8
Population	20	η	0.8	T	10	$P^{\text{Earthquake}}$	0.2746
Selection	0.1	ρ	36.13	$n^{\text{K,max}}$	4	$RR^{\text{Conventional}}$	0.97
		ζ	0.015	$T^{\text{tree}}_{\text{max}}$	3	$N^{\text{ASP,max}}$	9

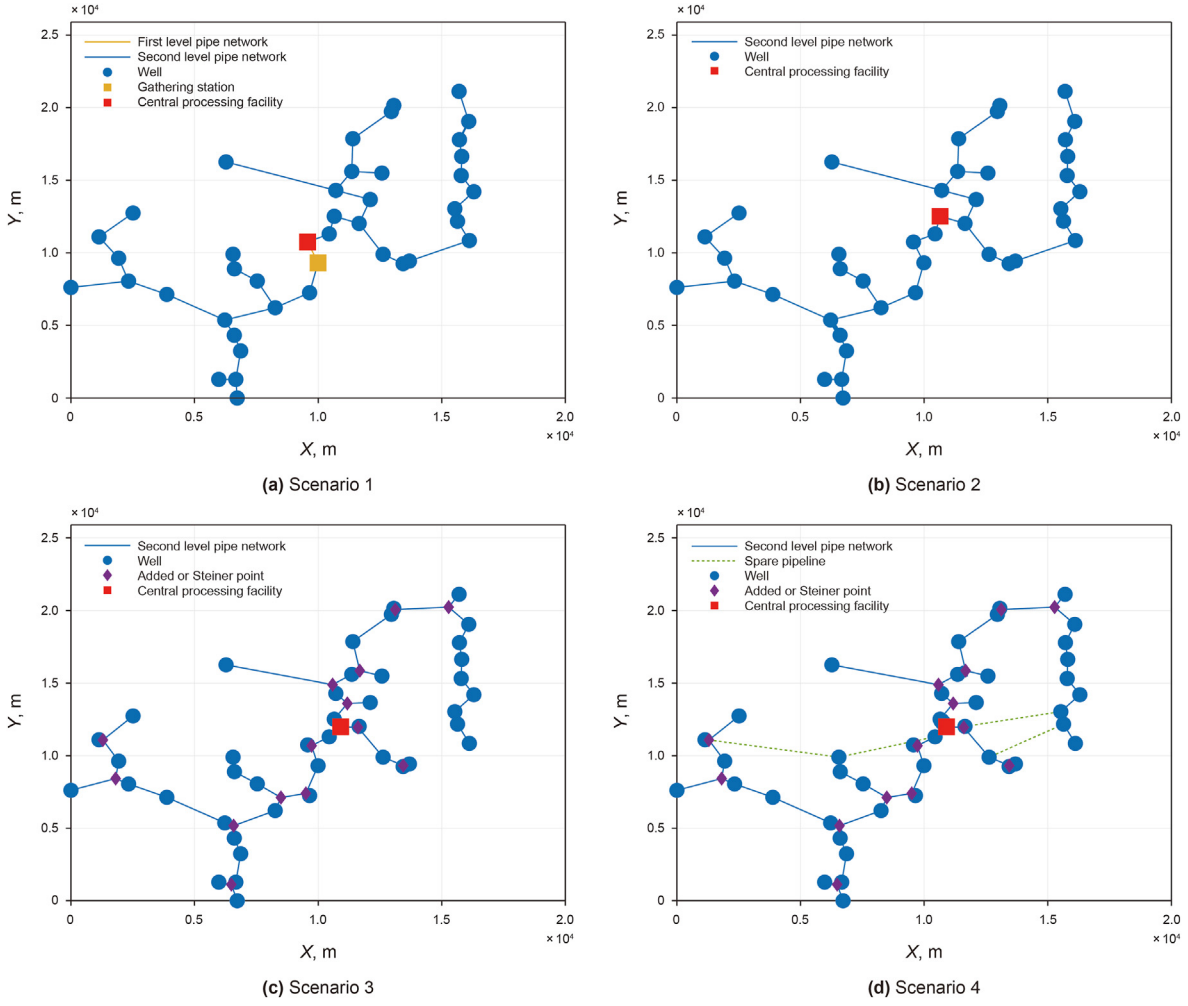

Fig. 6. Pipeline topology of scenarios 1 to 4.

Table 4
Cost results of scenarios 1 to 4.

Cost Name	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Pressure Loss Cost, CNY·a^{-1}	2.08×10^7	2.08×10^7	2.07×10^7	2.07×10^7
Pipeline Cost, CNY·a^{-1}	2.73×10^7	2.72×10^7	2.73×10^7	4.66×10^7
Station Cost, CNY·a^{-1}	1.42×10^7	8.24×10^6	8.24×10^6	8.24×10^6
Conventional Failure Cost, CNY·a^{-1}	6.04×10^7	5.81×10^7	5.21×10^7	1.14×10^7
Earthquake Failure Cost, CNY·a^{-1}	2.73×10^7	2.71×10^7	2.65×10^7	8.34×10^6
Total Annual Cost, CNY·a^{-1}	1.50×10^8	1.41×10^8	1.35×10^8	9.52×10^7
Pipeline Length, m	6.71×10^4	6.71×10^4	6.55×10^4	7.04×10^4
Conventional-Reliability	0.72	0.73	0.76	0.95
Earthquake-Reliability	0.67	0.67	0.68	0.90

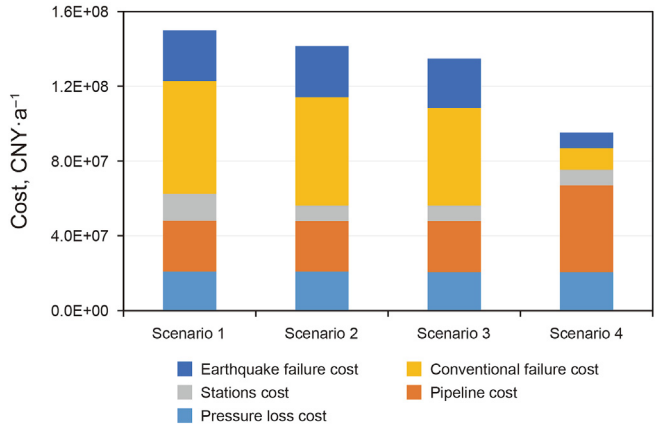

Fig. 7. Comparison of each cost of scenarios 1–4.

$$p_{(i,j)}^{\text{Magnitude}}(M) = 1170 \times \exp(-2.15 \times M), \quad \forall (i,j) \in \mathbf{E_T} \qquad (40)$$

The main parameters in the model are shown in Table 3.

5. Results and discussion

5.1. Results for scenarios 1 to 4

The calculated network topologies are shown in Fig. 6. Fig. 6a illustrates the initial topology of the gathering system of shale gas, while the optimized topologies of the remaining scenarios are presented in Fig. 6b–d. The results in Table 4 show the different cost-related items, namely pressure loss cost, pipeline cost, station cost, conventional failure cost, and earthquake failure cost.

5.2. Discussion of four scenarios

5.2.1. Economic discussion

The total cost and their corresponding contributions for all scenarios are shown in Fig. 7. As the number of optimization variables increases, the total annual cost tends to decrease, among them, the reduction of Scenario 4 is the most obvious. Specifically, the total annual cost of Scenario 4 is reduced by 5.48×10^7 CNY·a^{-1}compared with Scenario 1. Moreover, in Scenarios 1 to 3, the costs of conventional failure and earthquake failure account for the largest proportion of the total annual cost, and the sum of these two costs exceed 50% or even 60%. In Scenario 4, after improving the reliability of the pipe network, pipeline costs account for the largest proportion, reaching 48.95%.

In Scenario 2, the number of well clusters is reduced from 2 to 1. This reduces the cost of the gas gathering station while other costs do not vary significantly, hence reducing the total annual cost. Different from Scenario 2, the optimal pipe network topology of Scenario 3 is ESMT. Compared with MST, ESMT reduces the pipe length by introducing Steiner points. The shorter the pipe length, the lower the risk of damage and energy loss, and therefore the lower the total annual cost. This results in a 4.60% reduction in the total annual cost. In Scenario 4, the optimization of spare pipelines further reduces the total annual cost by 29.48%, compared with Scenario 3. On one hand, adding spare pipelines increases the investment cost and is thus unfavorable for reducing the total annual cost. On the other hand, the increase of spare pipelines provides new paths for gas transmission among well clusters, which reduces the risk of pipeline failures, thereby reducing conventional and earthquake failure. It can be concluded that there is a balance between the pipeline cost and the reliability of the pipeline network. In Scenario 4, due to the significant reduction in risk loss, the adverse impact of a small number of spare pipelines on the cost is completely offset, so the total annual cost is greatly reduced. The above results prove the need for an optimal cluster number, topology type, and reliability.

5.2.2. Reliability discussion

Fig. 8 reports the total annual cost, along with the trend of those reliabilities for conventional and earthquake scenarios. With the increase of the factors (K-value, pipeline topology and spare pipes) considered in the model, the conventional reliability and earthquake reliability of the pipeline network has been improved, and the total annual cost has been continuously reduced. Additionally, both conventional reliability and earthquake reliability have similar changing trends, and their influences on *TAC* are also consistent.

There is no significant variation between both reliability estimations in Scenario 2. In Scenario 3, the conventional reliability increases, whereas the earthquake reliability does not. The reason is that the total length of the ESMT pipeline network in Scenario 3 is shorter than in Scenarios 1–2, and the conventional reliability decreases as the length of the pipe network increases, as shown in Eqs. (25), (28)–(31). Furthermore, the earthquake reliability of the

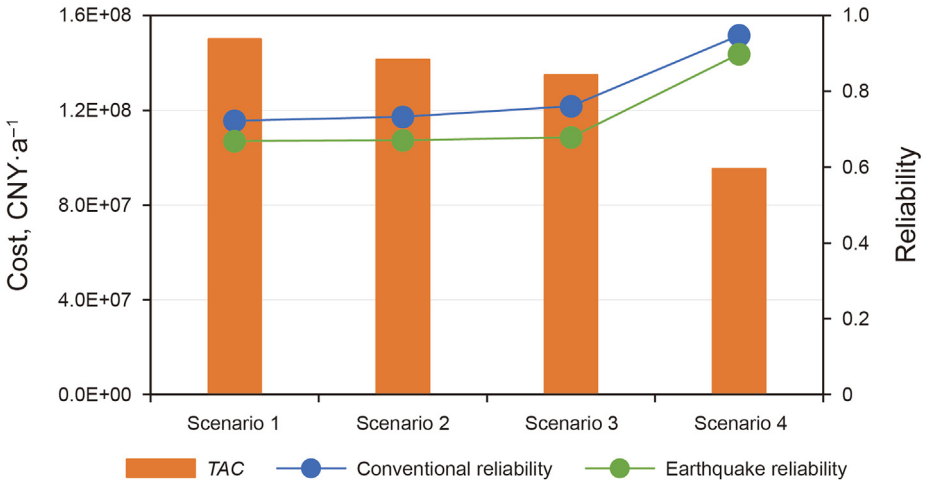

Fig. 8. The variations of TAC, conventional reliability, and earthquake reliability of scenarios 1–4.

Table 5
The conditions of sensitivity analysis.

Sensitivity analysis	K-value	Pipeline topology	Added spare pipeline
Well cluster number	Variable	First level: ESMT Second level: ESMT	0
Pipeline topology	1	First level: Variable Second level: Variable	0
Spare pipeline number	1	First level: ESMT Second level: ESMT	Variable

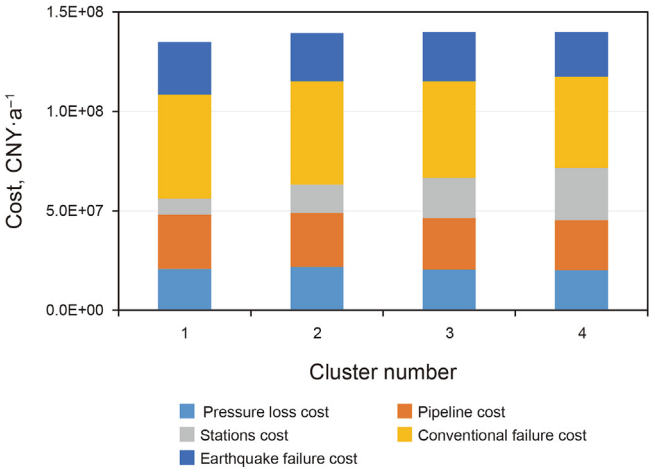

Fig. 9. The variations of each cost with cluster number.

Table 6
The variations of conventional and earthquake reliabilities with cluster number.

Cluster Number	1	2	3	4
Conventional reliability	0.760	0.761	0.778	0.789
Earthquake reliability	0.678	0.697	0.698	0.728

pipeline network is not only affected by the length, but also limited by the magnitude, depth, and epicenter distance of the earthquake. These earthquake factors increase the failure rate of a single pipeline, thus limiting the increase in earthquake reliability. In Scenario 4, both conventional reliability and earthquake reliability are greatly improved by 31.17% and 34.39% respectively. The increased reliability greatly reduces the total annual cost. The addition of spare pipelines directly modifies the topological structure of the pipe network, and new gathering paths are generated. Regardless of the conventional situation or the earthquake situation, new gathering paths means that even if the original route fails, the gathering of the shale gas will be still very likely to operate normally, so the reliability is significantly improved.

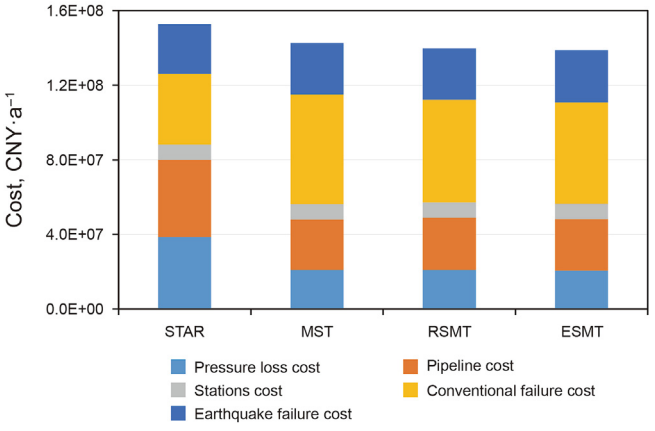

Fig. 10. The variations of each cost with pipeline topologies.

Based on the above results, we can infer that the cluster number has no apparent influence on the two reliabilities of the pipeline network. The topology type can further affect the conventional reliability by affecting the total length. Also, the spare pipeline can greatly improve the conventional reliability and the earthquake reliability at the same time.

5.3. Sensitivity analysis

The analysis in Section 5.2 is valid when the number of clusters, pipe network structure and the number of spare pipelines are fixed. However, these factors will vary in different real-life cases. In industrial production, designers are more concerned about the impact of changes in these factors on the economic benefits, and then determine the optimal layout of the case. Based on the results in Section 5.2, a sensitivity analysis for the cluster number, topology type, and the number of spare pipelines, on the economic benefit is carried out. The details of each sensitivity study are shown in Table 5.

5.3.1. Well cluster number

In this section, the range of clusters number is set between 1 and 4. Fig. 9 and Table 6 respectively show the costs and reliabilities for each of these different scenarios. The results in Fig. 9 further validate what has already been shown in Section 5.2. When the number of clusters is 1, the total annual cost is the lowest, given that the increase in cluster number will significantly increase the station cost. On the other hand, both the reliabilities of the pipe network tend to increase slightly as the number of clusters increases. However, the improvements of the reliabilities do not exceed 5%. Similarly, pipeline cost and pressure loss cost decrease slightly as the cluster number increases. Therefore, changes in station cost due to the number of clusters have the greatest impact on the objective function. To maximize the economic benefit, in this case, a lower number of clusters may be more appropriate. Overall, the optimal number of well clusters will vary according to actual engineering factors. If the locations of wells are not as concentrated as in this case study, but scattered in more areas and have longer distances, the costs related to the pipeline will be increased. In this case, increasing the number of clusters may reduce the total cost.

Table 7
The variations of conventional and earthquake reliability among pipeline topologies.

Topology	STAR	MST	RSMT	ESMT
Conventional Reliability	0.825	0.730	0.747	0.750
Earthquake Reliability	0.679	0.667	0.659	0.668

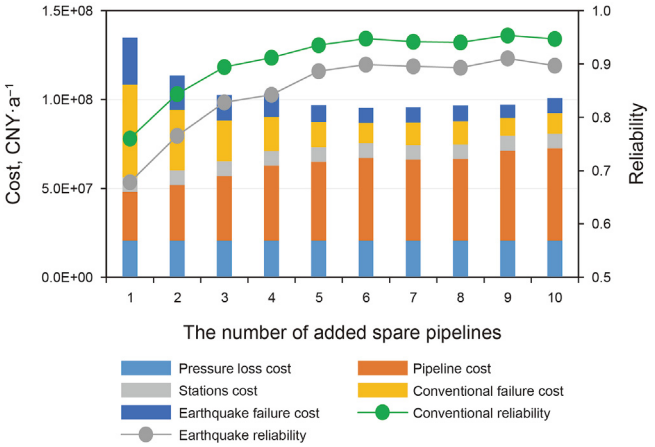

Fig. 11. The variations of each cost and reliability with spare pipeline number.

5.3.2. Pipeline topology

This section analyzes the four types of pipe network topologies, namely STAR, MST, RSMT, and ESMT, respectively. Fig. 10 shows the cost for these different topologies. The results prove one conclusion in Section 5.2 again, that is, the total annual cost of the pipe network with the ESMT structure is the lowest when other variables are kept fixed. In addition, changes in topology have a relatively larger impact on reliabilities and pipeline cost compared with cluster numbers. Thus, the consideration of pipeline topology in the optimization is necessary.

In the STAR pipeline network, each well is directly connected to the gas gathering station, and the pipelines are independent of each other. The damage of a pipeline will not affect the gas transportation of the wells that are not directly connected to it in the entire pipeline network. Therefore, the STAR pipeline network has higher reliability, as shown in Table 7. However, this leads to a longer length, which results in higher pipeline and pressure loss costs. Compared with the STAR pipe network, the MST pipe network connects each adjacent well and then connects a single well to the gas gathering station to achieve the gas gathering and transportation, reducing the total pipeline length, so its total annual cost is lower. By introducing Steiner points to shorten the length, the RSMT, and ESMT pipe networks further reduce the total annual cost and ESMT has the shortest length. Therefore, to obtain the maximum economic benefits, the ESMT pipe network is the better choice.

5.3.3. Spare pipeline number

This section sets the number of spare pipelines from 0 to 9. Fig. 11 shows the changes in different costs and reliabilities. The spare pipeline has a significant impact on the pipeline cost and both reliabilities. As the number of spare pipelines increases, the total annual cost and reliabilities show a decreasing and increasing trend, respectively. When the number of spare pipelines exceeds 5, the total annual cost rebounds slightly and the two reliability tend to stabilize. The main reasons for the change in the total annual cost come from the conventional and earthquake failures and pipeline costs. When the spare pipeline number is small, the proportion of failure costs is greater, and its decrease is the main reason for the decrease in the total annual cost. However, as the number of pipelines increases, the proportion of failure costs decreases, and the space for reduction is smaller. From this point onwards, pipeline costs dominate over the changes in the total annual cost, so the total annual cost tends to rebound. In the early stage of the increase in the number of pipelines, the increase in reliability was due to the appearance of a small number of spare pipelines that change the original network topology and provide new transportation paths. In the later stage, because the reliability of the pipeline network is at a high level, and alternative paths appear, the continuous increase of spare pipelines no longer has a significant impact on reliability. Overall, to maximize economic benefit, adding 5 spare pipelines seems to be a more suitable choice.

6. Conclusion

This research proposed a method to comprehensively optimize shale gas gathering systems considering reliability, uncertainty, topology arrangements, and hydraulic calculations. The optimization includes well clustering, selection of site and central processing stations, and pipe network topology. Additionally, the proposed model can describe the optimization considering several uncertainty factors. The uncertainty factors include pipeline failure under no-earthquake conditions, seismic location, depth, and magnitude. Using the proposed method, new topology arrangements with loops based on STAR, MST, RSMT and ESMT can be obtained, and the spare pipelines can also effectively enhance the reliability of the gathering system. The GeoSteiner algorithm is innovatively coupled with a set of optimization methods, namely genetic algorithm, K-means algorithm, Kruskal algorithm, and linear programming. A case study is carried out, where the integrated algorithm was able to provide realistic solutions and reaches the following insights:

(1) The shale gas gathering system with one well cluster, ESMT topology, and 5 spare pipelines has the lowest total annual cost with 9.52×10^7 CNY·a^{-1}. Its conventional reliability and earthquake reliability both reach 0.9.
(2) Well cluster number has a greater influence on station cost but has little impact on pipeline cost, pressure loss cost and failure costs. The fewer the cluster number, the lower the total annual cost and reliability of the pipe network system.
(3) Topology arrangements have a relatively larger impact on reliabilities and pipeline costs than cluster numbers. In general, a pipeline network with a tree-like topology is better suited to a star-like topology, although the former is less reliable. ESMT is the best among the three tree-like topologies.
(4) Spare pipeline number has the most significant impact on pipeline cost and reliability. With the increase in the number of spare pipelines, the total annual cost shows a decreasing trend at the beginning, followed by an increasing one later, while both reliabilities show an asymptotic behavior.

In future work, the proposed algorithm will be improved to solve the proposed problem on a larger scale more efficiently. Besides, more engineering factors are needed to be considered in the model, such as the topography of the well area, the variation of shale gas production with mining time, safety factors in the process of shale gas transportation, etc.

Acknowledgment

Financial supports from the National Natural Science Foundation of China under Grant (No. 22022816 and 22078358) are gratefully acknowledged.

Nomenclature

Main Abbreviations

CPF Central processing facility

ESMT	Euclidean Steiner minimum tree
GA	Genetic algorithm
GGS	Gathering gas station
LP	Linear programming
MILP	Mixed integer linear programming
MST	Minimum spanning tree
RSMT	Rectilinear Steiner minimal tree

Input Parameters

a	Coefficient of exponential distribution
a^{CPF}	Construction cost of the central processing station, CNY
$a^{\mathrm{Electricity}}$	Cost of electricity, CNY·kW^{-1}·h^{-1}
a_k^{GGS}	Cost of the construction of kth gathering gas station, CNY
$a^{\mathrm{Shale\ gas}}$	Unit price of shale gas, CNY·m^{-3}
b	Coefficient of exponential distribution
$b_{\alpha,\beta,\gamma}^{\mathrm{Binary}}$	Boolean parameter
C^{CPF}	Annual construction cost of the central processing station, CNY·a^{-1}
$c^{\mathrm{Earthquake}}$	Confidence level of reliability under earthquake conditions
C_k^{GGS}	Annual cost of the construction of kth gathering gas station, CNY·a^{-1}
I	Annual interest rate
$n^{\mathrm{ASP,max}}$	Max number of added spare pipelines
$n^{\mathrm{K,max}}$	Max number of well clusters
n^{s}	The number of Monte Carlo simulations for different parameters of earthquake
$P^{\mathrm{Earthquake}}$	The probability of the earthquake
$RR^{\mathrm{Conventional}}$	Unit reliability of the pipeline under conventional conditions, km^{-1}
T	Operation life cycle of the pipeline network system, a
t	Annual operating hours, h
$\mathrm{T}_{\mathrm{max}}^{\mathrm{tree}}$	Max number of the topology types of pipeline network
$u_{(i,j)}$	The flow rate of fluid in the pipeline, m·s^{-1}

Variables

$\mathrm{A}_{n^{V_T}\times n^{V_T}}^{(k)}$	Adjacent matrix
$a_{i,j}^{\mathrm{Conventional}}$	The element of the $\mathrm{A}_{n^{V_T}\times n^{V_T}}^{(k)}$
$a_{(i,j)}^{\mathrm{Pipeline}}$	The unit price of the pipeline $(i,\ j)$, CNY·m^{-1}
$a_{i,j}^{\mathrm{Earthquake}}$	The element of the $\mathrm{A}_{n^{V_T}\times n^{V_T}}^{(k)}$
$C_{(i,j)}^{Pipeline}$	Pipeline construction cost of the pipeline connected ith and jth vertex, CNY·a^{-1}
$C_{(i,j)}^{\mathrm{Pressure}}$	Pressure loss cost of the pipeline connected ith and jth vertex, CNY·a^{-1}
$C^{\mathrm{Reliability}}$	Loss cost of pipeline failure, CNY·a^{-1}
D	An exponent, which can control the flow direction in the pipeline
$D_{(i,j)}^{\mathrm{out}}$	The outer diameter of the pipeline connected ith and jth vertex, m
$D_{(i,j)}^{\mathrm{in}}$	The inner diameter of the pipeline connected ith and jth vertex, m
$G_{(i,j)}$	The mass flux of the fluid in the pipeline, kg·m^{-2}·s^{-1}
H	Depth of focus, km
$H_{(i,j)}^{\mathrm{f}}$	Head loss caused by resistance from ith vertex to jth vertex, m
$L_{(i,j)}$	The length of the pipeline connected ith and jth vertex, m
M	Magnitude of the earthquake
$\mathrm{M}_{n^{V_T}\times n^{V_T}}$	The reachability matrix
$m_{i,CPF}$	The element of $\mathrm{M}_{n^{V_T}\times n^{V_T}}$
n^{ASP}	The number of added spare pipelines
$N_{(i,j)}$	The shaft power that overcomes resistance when the gas is transported from ith vertex to jth vertex, W
n^{K}	The number of well clusters
n^{V_T}	The number of vertices in the optimized connected tree
$Ne_{(i,j)}$	The effective power of the conveyor between ith vertex and jth vertex, W
$PGV_{(i,j)}$	Peak ground velocity, cm·s^{-2}
$p_{(i,j)}^{\mathrm{Depth}}(H)$	The probability of seismic depth H when the earthquake occurs
$p_{(i,j)}^{\mathrm{Magnitude}}(M)$	The probability of seismic magnitude M when the earthquake occurs
$p_{(i,j)}^{\mathrm{in}}$	The inlet pressure of the pipeline, kPa
$p_{(i,j)}^{\mathrm{out}}$	The outlet pressure of the pipeline, kPa
Q_i	Shale gas production of ith well, kg·s^{-1}
$q_{l,i}$	The shale gas flowrate from ith vertex to the CPF in the lth simulation for the location of pipe failure under the conventional conditions, kg·s^{-1}
$q_{l,s,i}$	The shale gas flowrate from ith vertex to the CPF in the lth simulation for the location of earthquake and sth simulation for the seismic depth and magnitude under the earthquake condition, kg·s^{-1}
$R^{\mathrm{Conventional}}$	The reliability of the pipeline network system under conventional conditions
$R_{(i,j)}$	The distance from the edge $(i,\ j)$ to the epicenter, km
$\overline{R}^{\mathrm{Earthquake}}$	The reliability of the pipeline network system in the event of an earthquake
$RR_{(i,j)}^{\mathrm{Earthquake}}$	The repair rate of the branch pipeline $(i,\ j)$ in the event of an earthquake, km^{-1}
$T_{\mathrm{first}}^{\mathrm{tree}}$	Integer variable, the topology type of the first level pipeline network
$T_{\mathrm{second}}^{\mathrm{tree}}$	Integer variable, the topology type of the second level pipeline network
TAC	Total Annual Cost, CNY·a^{-1}
$v_{(i,j)}^{\mathrm{scaled}}$	Scaled velocity on the edge $(i,\ j)$
$w_{(i,j)}$	The mass flow rate of the fluid in the pipeline, kg·s^{-1}
W_γ	Mass flow rate of vertex γ , kg·s^{-1}
$w_{\alpha,\beta,\gamma}$	the mass flow rate of shale gas in the edge $(\alpha,\ \beta)$ connected to vertex γ when the material balance is performed on the vertex γ, kg·s^{-1}
$Wt_{(i,j)}$	The weight of the unit length of the pipeline, kg·s^{-1}

Sets and indices

$\mathbf{E_T}$	The set of edges in the connected graph, denoted by indices $(i,\ j)$
$\mathbf{V_T}$	The set of vertices in the connected graph, denoted by indices i,j,α,β,γ
$\mathbf{V^{GGS}}$	The set of gas gathering stations, denoted by index k
$\mathbf{N^S}$	The set of seismic depth and magnitude simulations, denoted by index s
$\mathbf{N^L}$	The set of pipeline failure location and earthquake location simulations, denoted by index l

Greek letters

η	The efficiency of conveying equipment
μ	The location parameter of the normal distribution
ρ	The density of the fluid in the pipeline, kg·m^{-3}
σ	The shape parameter of the normal distribution
ζ	The coefficient of friction of the pipeline

References

Adachi, T., Ellingwood, B.R., 2008. Serviceability of earthquake-damaged water systems: effects of electrical power availability and power backup systems on system vulnerability. Reliab. Eng. Syst. Saf. 93 (1), 78–88. https://doi.org/

10.1016/j.ress.2006.10.014.
Alsharqawi, M., Zayed, T., Parvizsedghy, L., et al., 2020. Reliability assessment model for water distribution networks. J. Pipeline Syst. Eng. Pract. 11 (2), 04019059. https://doi.org/10.1061/(ASCE)PS.1949-1204.0000442.
Arredondo-Ramirez, K., Ponce-Ortega, J.M., El-Halwagi, M.M., 2016. Optimal planning and infrastructure development for shale gas production. Energy Convers. Manag. 119, 91–100. https://doi.org/10.1016/j.enconman.2016.04.038.
Baeza, D., Ihle, C.F., Ortiz, J.M., 2017. A comparison between ACO and Dijkstra algorithms for optimal ore concentrate pipeline routing. J. Clean. Prod. 144, 149–160. https://doi.org/10.1016/j.jclepro.2016.12.084.
Cariou, P., Cheaitou, A., Larbi, R., et al., 2018. Liner Shipping Network Design with Emission Control Areas: A Genetic Algorithm-Based Approach, vol. 63. Transportation Research Part D: Transport and Environment, pp. 604–621. https://doi.org/10.1016/j.trd.2018.06.020.
Cen, K., Zheng, X., Jiang, X., et al., 2016. A three-level optimization methodology for the partitioning of shale gas wellpad groups. J. Nat. Gas Sci. Eng. 34, 341–355. https://doi.org/10.1016/j.jngse.2016.07.009.
Chan, A.L.S., Hanby, V.I., Chow, T.T., 2007. Optimization of distribution piping network in district cooling system using genetic algorithm with local search. Energy Convers. Manag. 48 (10), 2622–2629. https://doi.org/10.1016/j.enconman.2007.05.008.
Chen, Y.Z., Cheng, X., Li, J., et al., 2020. A multi-level programming for shale gas-water supply chains accounting for tradeoffs between economic and environmental concerns. Comput. Chem. Eng. 135, 106761. https://doi.org/10.1016/j.compchemeng.2020.106761.
Chen, Y.Z., He, L., Li, J., et al., 2018. Multi-criteria design of shale-gas-water supply chains and production systems towards optimal life cycle economics and greenhouse gas emissions under uncertainty. Comput. Chem. Eng. 109, 216–235. https://doi.org/10.1016/j.compchemeng.2017.11.014.
Christodoulou, S.E., Fragiadakis, M., Agathokleous, A., et al., 2018. In: Chapter 5 - Vulnerability Assessment of Water Distribution Networks under Seismic Loads. Urban Water Distribution Networks. Butterworth-Heinemann, pp. 173–207. https://doi.org/10.1016/B978-0-12-813652-2.00005-0. S. E. Christodoulou, M. Fragiadakis, A. Agathokleous and S. Xanthos.
Dementiev, R., Sanders, P., Schultes, D., et al., 2004. Engineering an external memory minimum spanning tree algorithm. In: 3rd IFIP International Conference on Theoretical Computer Science. Kluwer, Toulouse, France. https://doi.org/10.1007/1-4020-8141-3_17.
Eia, U.S., DoE, U.S., 2013. Technically Recoverable Shale Oil and Shale Gas Resources: an Assessment of 137 Shale Formations in 41 Countries outside the United States, vol. 1. US Energy Information Agency & Advanced Research Institute; U.S. Department of Energy, Washington, USA, pp. 1–10. https://www.eia.gov/analysis/studies/worldshalegas/pdf/overview.pdf.
Grossmann, I.E., Apap, R.M., Calfa, B.A., et al., 2016. Recent advances in mathematical programming techniques for the optimization of process systems under uncertainty. Comput. Chem. Eng. 91, 3–14. https://doi.org/10.1016/j.compchemeng.2016.03.002.
Hong, B.Y., Li, X.P., Di, G.J., et al., 2019. An integrated MILP method for gathering pipeline networks considering hydraulic characteristics. Chem. Eng. Res. Des. 152, 320–335. https://doi.org/10.1016/j.cherd.2019.08.013.
Hong, C., Estefen, S.F., Wang, Y.X., et al., 2018. An integrated optimization model for the layout design of a subsea production system. Appl. Ocean Res. 77, 1–13. https://doi.org/10.1016/j.apor.2018.05.009.
Hosseini, M., MirHassani, S.A., 2015. Refueling-station location problem under uncertainty. Transport. Res. E Logist. Transport. Rev. 84, 101–116. https://doi.org/10.1016/j.tre.2015.10.009.
Juhl, D., Warme, D.M., Winter, P., et al., 2018. The GeoSteiner software package for computing Steiner trees in the plane: an updated computational study. Mathematical Programming Computation 10 (4), 487–532. https://doi.org/10.1007/s12532-018-0135-8.
Kerzmann, T.L., Buxton, G.A., Preisser, J., 2014. A computer model for optimizing the location of natural gas fueling stations. Sustain. Energy Technol. Assessments 7, 221–226. https://doi.org/10.1016/j.seta.2013.10.004.
Kim, Y., Kang, W.H., 2013. Network reliability analysis of complex systems using a non-simulation-based method. Reliab. Eng. Syst. Saf. 110, 80–88. https://doi.org/10.1016/j.ress.2012.09.012.
Larsson, C., 2018. Chapter 6 - Clustering. 5G Networks. C. Larsson. Academic Press, pp. 123–141. https://doi.org/10.1016/B978-0-12-812707-0.00011-5.
Li, F.F., Liu, Q., Guo, X., et al., 2015. A survey of optimization method for oil-gas pipeline network layout. International Conference on Mechatronics, Electronic, Industrial and Control Engineering (MEIC). Shenyang, China 8, 257–260 (in Chinese).
Li, H.M., Xia, Q.Y., Wang, Y., 2017. Research and improvement of Kruskal algorithm. J. Comput. Commun. 5 (12), 63–39 (in Chinese).
Li, Z.K., Ierapetritou, M., 2008. Process scheduling under uncertainty: review and challenges. Comput. Chem. Eng. 32 (4–5), 715–727. https://doi.org/10.1016/j.compchemeng.2007.03.001.
Lin, R.H., Ye, Z.Z., Guo, Z.Y., et al., 2020. Hydrogen station location optimization based on multiple data sources. Int. J. Hydrogen Energy 45 (17), 10270–10279. https://doi.org/10.1016/j.ijhydene.2019.10.069.
Liu, Z., Pan, B., 2003a. Optimization of submarine oil and gas gathering and transportation pipeline system under reliability constraints. Ocean Eng. 21 (2), 26–31. https://doi.org/10.16483/j.issn.1005-9865.2003.02.005 (in Chinese).
Liu, Z., Pan, B., 2003b. Reliability evaluation for subsea oil-gas pipeline networks. Ocean Eng. 21 (4), 104–110. https://doi.org/10.1007/s11769-003-0089-1 (in Chinese).
Meng, Z., Guo, L.B., Hao, P., et al., 2021a. On the use of probabilistic and non-probabilistic super parametric hybrid models for time-variant reliability analysis. Comput. Methods Appl. Mech. Eng. 386, 114113. https://doi.org/10.1016/j.cma.2021.114113.
Meng, Z., Li, G., Wang, B.P., et al., 2015. A hybrid chaos control approach of the performance measure functions for reliability-based design optimization. Comput. Struct. 146, 32–43. https://doi.org/10.1016/j.compstruc.2014.08.011.
Meng, Z., Li, G., Wang, X., et al., 2021b. A comparative study of metaheuristic algorithms for reliability-based design optimization problems. Arch. Comput. Methods Eng. 28 (3), 1853–1869. https://doi.org/10.1007/s11831-020-09443-z.
Meng, Z., Li, G., Yang, D.X., et al., 2017. A new directional stability transformation method of chaos control for first order reliability analysis. Struct. Multidiscip. Optim. 55 (2), 601–612. https://doi.org/10.1007/s00158-016-1525-z.
Nemati, M., Braun, M., Tenbohlen, S., 2018. Optimization of unit commitment and economic dispatch in microgrids based on genetic algorithm and mixed integer linear programming. Appl. Energy 210, 944–963. https://doi.org/10.1016/j.apenergy.2017.07.007.
Oke, D., Mukherjee, R., Sengupta, D., et al., 2020. On the optimization of water-energy nexus in shale gas network under price uncertainties. Energy 203, 117770. https://doi.org/10.1016/j.energy.2020.117770.
Qiang, L., Li, M., Li, F.F., 2016. An intelligent optimization method for oil-gas gathering and transportation pipeline network layout. In: 28th Chinese Control and Decision Conference (CCDC). Yinchuan, China, pp. 4621–4626 (in Chinese).
Roetzel, W., Luo, X., Chen, D.Z., 2020. In: Chapter 6 - Optimal Design of Heat Exchanger Networks. Design and Operation of Heat Exchangers and Their Networks. Academic Press, pp. 231–317. https://doi.org/10.1016/B978-0-12-817894-2.00006-6. W. Roetzel, X. Luo and D. Chen.
Sanaye, S., Mahmoudimehr, J., 2013. Optimal design of a natural gas transmission network layout. Chem. Eng. Res. Des. 91 (12), 2465–2476. https://doi.org/10.1016/j.cherd.2013.04.005.
Stijepovic, M.Z., Linke, P., 2011. Optimal waste heat recovery and reuse in industrial zones. Energy 36 (7), 4019–4031. https://doi.org/10.1016/j.energy.2011.04.048.
Su, H., Zio, E., Zhang, J.J., et al., 2019. A method for the multi-objective optimization of the operation of natural gas pipeline networks considering supply reliability and operation efficiency. Comput. Chem. Eng. 131, 106584. https://doi.org/10.1016/j.compchemeng.2019.106584.
Takagi, D., Koyama, K., Ishizuka, A., et al., 2002. Reliability analysis system for risk management of LNG receiving terminals and piping network. In: Annual Reliability and Maintainability Symposium. 2002 Proceedings (Cat. No.02CH37318). https://doi.org/10.1109/RAMS.2002.981643.
Tromans, I.J., 2004. Behaviour of Buried Water Supply Pipelines in Earthquake Zones. Ph.D., Imperial College London.
Wang, S.Q., 2018. Shale gas exploitation: status, problems and prospect. Nat. Gas. Ind. B 5 (1), 60–74. https://doi.org/10.1016/j.ngib.2017.12.004.
Warme, D.M., 1998. Spanning Trees in Hypergraphs with Applications to Steiner Trees. Ph.D., University of Virginia.
Wu, J.X., Wang, P.F., 2021. Post-disruption performance recovery to enhance resilience of interconnected network systems. Sustainable and Resilient Infrastructure 6 (1–2), 107–123. https://doi.org/10.1080/23789689.2019.1710073.
Wu, Y., Xia, T.Q., Jatowt, A., et al., 2021. Context-aware heatstroke relief station placement and route optimization for large outdoor events. Int. J. Health Geogr. 20 (1), 23. https://doi.org/10.21203/rs.3.rs-219494/v1.
Wu, Y., Zhang, S., Wang, R.Q., et al., 2020a. A design methodology for wind farm layout considering cable routing and economic benefit based on genetic algorithm and GeoSteiner. Renew. Energy 146, 687–698. https://doi.org/10.1016/j.renene.2019.07.002.
Wu, Y., Zhang, S., Wang, R.Q., et al., 2020b. New model for large scale chemical industrial layout optimization. Chem. Eng. Res. Des. 161, 58–71. https://doi.org/10.1016/j.cherd.2020.06.026.
Yoon, S., Lee, Y.J., Jung, H.J., 2018. A comprehensive framework for seismic risk assessment of urban water transmission networks. Int. J. Disaster Risk Reduc. 31, 983–994. https://doi.org/10.1016/j.ijdrr.2018.09.002.
Yu, W.C., Gong, J., Song, S.F., et al., 2019. Gas supply reliability analysis of a natural gas pipeline system considering the effects of underground gas storages. Appl. Energy 252, 113418. https://doi.org/10.1016/j.apenergy.2019.113418.
Yu, W.C., Wen, K., Min, Y., et al., 2018. A methodology to quantify the gas supply capacity of natural gas transmission pipeline system using reliability theory. Reliab. Eng. Syst. Saf. 175, 128–141. https://doi.org/10.1016/j.ress.2018.03.007.
Zachariasen, M., 1998. Rectilinear full steiner tree generation. Networks 33 (2), 125–143. https://doi.org/10.1002/(SICI)1097-0037 (199903)33:2<125::AID-NET4>3.0.CO;2-S.
Zhang, H.R., Liang, Y.T., Ma, J., et al., 2017a. An improved PSO method for optimal design of subsea oil pipelines. Ocean Eng. 141, 154–163. https://doi.org/10.1016/j.oceaneng.2017.06.023.

Zhang, H.R., Liang, Y.T., Zhang, W., et al., 2017b. A unified MILP model for topological structure of production well gathering pipeline network. J. Petrol. Sci. Eng. 152, 284–293. https://doi.org/10.1016/j.petrol.2017.03.016.

Zhou, J., Gong, J., Li, X., et al., 2014a. Optimization of coalbed methane gathering system in China. Adv. Mech. Eng. 6, 147381. https://doi.org/10.1155/2014/147381.

Zhou, J., Li, X., Cheng, M., et al., 2014b. Topology analysis and optimization design of coal bed methane gathering system in China. In: 10th International Pipeline Conference, vol. 4, 33278. https://doi.org/10.1115/IPC2014-33278. Calgary, Alberta, Canada.

Zhou, J., Peng, J.H., Liang, G.C., et al., 2019. Layout optimization of tree-tree gas pipeline network. J. Petrol. Sci. Eng. 173, 666–680. https://doi.org/10.1016/j.petrol.2018.10.067.