Vol. 34, No. 1 (2024) | DOI: 10.37190/ord240108

OPEN ACCESS

Operations Research and Decisions

www.ord.pwr.edu.pl

OPERATIONS RESEARCH AND DECISIONS QUARTERLY

A solution method for stochastic multilevel programming problems. A systematic sampling evolutionary approach

Natnael Nigussie Goshu[1*] Semu Mitiku Kassa[2]

[1]Mathematics Department, Addis Ababa Science and Technology University, Addis Ababa, Ethiopia

[2]Department of Mathematics and Statistical Sciences, Botswana International University of Science and Technology, Pvt Bag 16, Palapye, Botswana

[*]Corresponding author, email address: natnael.nigussie@aastu.edu.et

Abstract

Stochastic multilevel programming is a mathematical programming problem with some given number of hierarchical levels of decentralized decision makers and having some kind of randomness properties in the problem definition. The introduction of some randomness property in its hierarchical structure makes stochastic multilevel problems computationally challenging and expensive. In this article, a systematic sampling evolutionary method is adapted to solve the problem. The solution procedure is based on realization of the random variables and systematic partitioning of each hierarchical level's decision space for searching an optimal reaction. The search goes sequentially upwards starting from the bottom up through the top hierarchical level problem. The existence of solution and convergence of the solution procedure is shown. The solution procedure is implemented and tested on some selected deterministic test problems from literature. Moreover, the proposed algorithm can be used to solve stochastic multilevel programming problems with additional complexity in their problem definition.

Keywords: *multilevel programming, stochastic programming, Stackelberg equilibrium, sample average approximation, systematic sampling, particle swarm optimization*

1. Introduction

Multilevel programming was first introduced in the field of game theory by Stackelberg in 1934 as a two-person leader-follower game [53]. Bracken and McGill studied it as a generalization of mathematical programming [17, 18]. Multilevel programming is defined as a sequential decision-making problem in a non-cooperative, decentralized, and multiple-level hierarchical procedure. The main task in such problems is to optimize the objective of the leader (upper level), which is constrained by the optimal choice of some variables by the followers, in a sequential structure. However, this kind of problem is usually hard

Received 21 January 2023, accepted 12 January 2024, published online 28 March 2024
ISSN 2391-6060 (Online)/© 2024 Authors

The costs of publishing this issue have been co-financed by the program *Development of Academic Journals* of the Polish Ministry of Education and Science under agreement RCN/SP/0241/2021/1

to solve and is categorized as NP-hard even for its simplest case the linear bilevel programming problem [11, 56]. The inclusion of non-linearity, non-convexity, non-differentiability, and other undesirable properties add further complexity to solving multilevel problems.

Proposed methods for solving multilevel programming problems include vertex enumeration [15, 51, 64] for solving linear bilevel and trilevel programming problems, sub-gradient or bundle method [52], reformulation method using the Karush-Kuhn-Tucker (KKT) conditions for solving bilevel programming [10], reformulation technique with Taylor's approximation for solving quad-level programming [30], optimal value reformulation technique for trilevel linear problems [8, 40], parametric approach [6, 7, 24, 32, 33, 45], data-driven optimization approach [12–14] and also fuzzy approach [37, 50].

Evolutionary algorithms have been presented for solving bilevel programming problems by converting them into single-level mathematical programming problems [3, 55, 57] and other special methods are also described in [22]. Reformulation technique with particle swarm optimization (PSO) is also proposed for solving trilevel problems [28, 29]. However, solutions that are obtained through the fuzzy goal programming method for multilevel problems are shown to be suboptimal [21].

A multi-parametric programming method is also used to solve multilevel programming problems and depending on the type of functions involved and the constraint set, it can have an analytic (exact) solution or approximated (non-exact) solution. For the case of bilevel linear [16], bilevel quadratic [1, 2], bilevel mixed integer linear programs [9] and bilevel quadratically constrained quadratic programs [42], it has been shown that analytic solutions can be obtained using the multi-parametric approach. In addition to this, the approach can also be naturally extended for solving the Stackelberg–Nash equilibrium type problems and for trilevel problems with polyhedral constraints [34, 45] and with multiple followers [35]. Even if the multi-parametric programming method can be extended to solve multilevel problems with any finite hierarchical levels, it works mainly for polyhedral-constrained problems. The recently proposed method in [62] may help to extend the multi-parametric programming approach to also solve problems with non-polyhedral constraints.

Specific solution approaches for solving general multilevel problems for any number of levels are proposed in [34, 54, 60]. In [54], evolutionary strategy is implemented for solving each decision-makers problem at each hierarchical level sequentially by fixing their strategy down through the bottom level problem and repeating this continuously until a termination criterion is fulfilled. Whereas, in [60] the authors assume that systematically selected decision variable values are sent by the leader to check the optimal reactions of the followers before deciding over its optimal decision variable value and continue this until a termination criterion is fulfilled. In addition to this, in [34] branch and bound multi-parametric method is proposed for smooth multilevel programming problems with polyhedral constraints that satisfy strict second-order complementarity conditions. In general, all those proposed methods can only be applied to the deterministic version of multilevel programming problems.

If randomness is introduced in a multilevel programming problem, we shall call such a problem a stochastic multilevel programming (SMLP) problem. SMLP is a generalization of a multilevel programming problem when the variance of the randomness in a probability distribution is different from zero. Basic principles of SMLP have been designed by Patricsson and Wynter [44] for solving structural programming problems. Christiansen [19] also formulated a topological optimization model in structural mechanics. There are various application problems listed in the open literature that use the idea of stochastic multilevel programming procedures in one way or the other. Namely, in transportation [4], in traffic modelling [43], in telecommunications problems [5, 58], in resource allocation problems [36], and the references in [25].

SMLP is classified into two categories. The first one is a two-stage stochastic multilevel program (TS-SMLP), where the decision system allows the decision maker at each hierarchical level to adjust his/her decisions when some randomness appears over time. The other category is a chance constraint multilevel program (CC-MLP), where the decision maker at each hierarchical level needs to look for fault tolerance and system reliability. In either of the cases, solving such kinds of problems is computationally expensive and challenging due to its hierarchical structure and the existence of randomness property.

Some of the proposed methods are reformulation methods with a scenario analysis approach for solving a two-stage stochastic bilevel programming problem [4, 58]. In those methods, the lower-level problem is assumed to be strictly convex and regular, and also continuous differentiability of all functions that are involved in the problem is required. Then, the program is reformulated into a single-level, two-stage stochastic program using KKT conditions as their solution technique. Similar reformulation technique is implemented for solving stochastic bilevel problems also in [23, 31, 61, 63]. However, all those proposed methods work only for two-level problems and cannot be extended to three or more levels of hierarchy.

For chance constraint type problems, a fuzzy goal programming approach is proposed by assigning some aspiration level [39, 46–48]. But, those proposed methods work only for problems in which randomness is involved only in the constraints, especially in the form of linear constraints, not in the objective functions. Moreover, the proposed optimization procedure in these methods may produce a suboptimal solution as it was indicated in [21]. Generally, a comprehensive method for solving SMLP problems is lacking, i.e., there are limitations in the study of the recourse as well as the chance constraint version of the problem for any number of hierarchical levels.

The main objective of this study is to develop a strategy for solving SMLP by extending the Systematic Sampling Evolutionary (SSE) method which was proposed by the authors for solving stochastic bilevel programming (SBLP) in [25]. The same method is further extended in [26] for bilevel problems with multiple followers cases, and in [27] for various forms of supply chain management problems. In the case of SBLP, the follower's problem becomes a global optimization problem or Nash equilibrium problem when the leader sets his/her strategy. Then any optimization technique can be implemented for solving the second level problem for a single follower case and Nash equilibrium solution technique for multiple followers case. However, in the case of SMLP having more than two levels, only the bottom levels can be a global optimization problem when higher-level hierarchical decision-makers set their strategy. Now, the difficulty in solving the SMLP problem is how to formulate strategies for the intermediate decision makers since each decision maker is assumed to have an infinite number of possible strategies. These infinitely induced strategies at each intermediate hierarchical level make any given solution method too expensive and challenging. So, some fundamental adjustments are required to be made to the SSE method to manage the strategies of the intermediate decision makers. The remaining part of the paper is structured as follows. In section 2, a general introduction to SMLP is discussed. In Section 3, solution procedures are presented. Simulation results are presented in section 4. An section 5, some remarks on SMLP are made.

2. Stochastic multilevel programming

Consider a probability space $(\Omega, \mathcal{F}, P)$, where Ω is a non-empty sample set, $\mathcal{F}$ is a σ-algebra of measurable subsets (events) of Ω and P is a probability measure on $\mathcal{F}$. Stochastic programming problems are mainly classified into two-stage stochastic programming and chance-constrained programming problems. In the next paragraph, a two-stage stochastic multilevel programming problem is presented and the formulation for the chance-constrained multilevel programming problem is presented at the end of Section 3.

A mathematical formulation for a two-stage stochastic bilevel programming is given by

$$
\begin{aligned}
&\min_{x_1} F_1(x_1,\, x_2) + \mathbb{E}_\omega[Q(x_1,\, x_2, \omega)] \\
&\text{subject to } G_1(x_1,\, x_2) \le 0, \text{ where } x_2 \text{ solves the problem} \\
&\qquad \min_{x_2} F_2(x_1,\, x_2) \\
&\qquad \text{subject to } G_2(x_1,\, x_2) \le 0,
\end{aligned} \tag{1}
$$

where, $\forall \omega \in \Omega$,

$$
Q(x_1,\, x_2, \omega) = \begin{cases}
\min_{\bar{x}_1} \bar{F}_1(\bar{x}_1(\omega), \bar{x}_2(\omega), \omega) \\
\text{subject to } \bar{G}_1(x_1, \bar{x}_1(\omega), x_2, \bar{x}_2(\omega), \omega) \le 0 \\
\qquad \text{where } \bar{x}_2(\omega) \text{ solves the problem} \\
\qquad \min_{\bar{x}_2} \bar{F}_2(\bar{x}_1(\omega), \bar{x}_2(\omega), \omega) \\
\qquad \text{subject to } \bar{G}_2(x_1, \bar{x}_1(\omega), x_2, \bar{x}_2(\omega), \omega) \le 0
\end{cases}
$$

with x_1 and x_2 are first-stage decision variables which are supposed to be decided before observing the random outcome at the second stage; $\bar{x}_1$ and $\bar{x}_2$ are second-stage decision variables which are supposed to be decided after all randomness properties have been removed.

In problem (1), assume that $Q(x_1,\, x_2, \omega)$, $\bar{F}_1(\bar{x}_1(\omega), \bar{x}_2(\omega), \omega)$ and $\bar{F}_2(\bar{x}_1(\omega), \bar{x}_2(\omega), \omega)$ are measurable functions with respect to ω. So, $\forall \omega \in \Omega$ problem (1) can be reformulated as a standard stochastic bilevel program in the following form [4]

$$
\min_{x_1, \bar{x}_1} F_1(x_1,\, x_2) + \mathbb{E}_\omega[\bar{F}_1(\bar{x}_1(\omega), \bar{x}_2(\omega), \omega)] \tag{2a}
$$
$$
\begin{aligned}
&\text{subject to } G_1(x_1,\, x_2) \le 0 \\
&\qquad \bar{G}_1(x_1, \bar{x}_1(\omega), x_2, \bar{x}_2(\omega), \omega) \le 0 \\
&\qquad \text{where } x_2 \text{ and } \bar{x}_2 \text{ solve the problem}
\end{aligned}
$$
$$
\min_{x_2, \bar{x}_2} F_2(x_1,\, x_2) + \mathbb{E}_\omega[\bar{F}_2(\bar{x}_1(\omega), \bar{x}_2(\omega), \omega)] \tag{2b}
$$
$$
\begin{aligned}
&\text{subject to } G_2(x_1,\, x_2) \le 0 \\
&\qquad \bar{G}_2(x_1, \bar{x}_1(\omega), x_2, \bar{x}_2(\omega), \omega) \le 0 \\
&\qquad x_i \in X_i, \bar{x}_i \text{ is a measurable function from } \Omega \text{ to } \bar{X}_i
\end{aligned}
$$

In general, for an ℓ level two-stage hierarchical decision problem, suppose that the vectors x_i, $1 \leq i \leq \ell$ represent the first stage decision variables, which are called here-and-now variables, and the vectors $\bar{x}_i$, $1 \leq i \leq \ell$ represent the second stage decision variables, which are called the wait-and-see variables. The here-and-now variables are assumed to be decided before observing the random outcome at the second stage, whereas the wait-and-see variables are supposed to be decided after all randomness properties have been removed. Then, such an ℓ level TS-SMLP problem with randomness in the given probability distribution can be given $\forall \omega \in \Omega$ by

$$
\begin{aligned}
&\min_{x_1,\bar{x}_1} F_1(x_1,\, x_2,\, \ldots,\, x_\ell) + \mathbb{E}_\omega[\bar{F}_1(\bar{x}_1(\omega), \bar{x}_2(\omega),\, \ldots,\, \bar{x}_\ell(\omega), \omega)] \qquad \text{(3a)}\\
&\text{subject to } G_1(x_1,\, x_2,\, \ldots,\, x_\ell) \leq 0\\
&\qquad \bar{G}_1(x_1, \bar{x}_1(\omega), x_2, \bar{x}_2(\omega),\, \ldots,\, x_\ell, \bar{x}_\ell(\omega), \omega) \leq 0,\\
&\quad \text{where } x_2, \bar{x}_2,\, \ldots,\, x_\ell \text{ and } \bar{x}_\ell \text{ solve the problem}\\
&\quad \min_{x_2,\bar{x}_2} F_2(x_1,\, x_2,\, \ldots,\, x_\ell) + \mathbb{E}_\omega[\bar{F}_2(\bar{x}_1(\omega), \bar{x}_2(\omega),\, \ldots,\, \bar{x}_\ell(\omega), \omega)]\\
&\quad \text{subject to } G_2(x_1,\, x_2,\, \ldots,\, x_\ell) \leq 0\\
&\qquad \bar{G}_2(x_1, \bar{x}_1(\omega), x_2, \bar{x}_2(\omega),\, \ldots,\, x_\ell, \bar{x}_\ell(\omega), \omega) \leq 0,\\
&\qquad \text{where } x_3, \bar{x}_3,\, \ldots,\, x_\ell \text{ and } \bar{x}_\ell \text{ solve the problem}\\
&\qquad\qquad\qquad \ddots\\
&\qquad\quad \text{where } x_\ell \text{ and } \bar{x}_\ell \text{ solve the problem}\\
&\qquad\quad \min_{x_\ell,\bar{x}_\ell} F_\ell(x_1,\, x_2,\, \ldots,\, x_\ell) + \mathbb{E}_\omega[\bar{F}_\ell(\bar{x}_1(\omega), \bar{x}_2(\omega),\, \ldots,\, \bar{x}_\ell(\omega), \omega)] \qquad \text{(3b)}\\
&\qquad\quad \text{subject to } G_\ell(x_1,\, x_2,\, \ldots,\, x_\ell) \leq 0\\
&\qquad\qquad \bar{G}_\ell(x_1, \bar{x}_1(\omega), x_2, \bar{x}_2(\omega),\, \ldots,\, x_\ell, \bar{x}_\ell(\omega), \omega) \leq,\, \ldots,\\
&\qquad\qquad x_i \in X_i, \bar{x}_i \text{ is a measurable function from } \Omega \text{ to } \bar{X}_i
\end{aligned}
$$

Here we study problem (3) having a decision variable space $\mathbb{R}^m$, such that $(x_1, \bar{x}_1, x_2, \bar{x}_2,\, \ldots,\, x_\ell, \bar{x}_\ell) \in \prod_{t=1}^{\ell}(X_t \times \bar{X}_t) =: \mathbb{S}(\neq \emptyset) \subseteq \mathbb{R}^m$, where $X_i = [l_i, u_i]^{m_i} \subseteq \mathbb{R}^{m_i}$. Moreover, for each hierarchical level $i \in \{1, 2,\, \ldots,\, \ell\}$, let $\mathbb{S}_i$ denote the set given by

$$
\begin{aligned}
\mathbb{S}_i =& \{(x_1, \bar{x}_1, x_2, \bar{x}_2,\, \ldots,\, x_\ell, \bar{x}_\ell) : G_i(x_1,\, x_2,\, \ldots,\, x_\ell)\\
&\leq 0, \bar{G}_i(x_1, \bar{x}_1(\omega), x_2, \bar{x}_2(\omega),\, \ldots,\, x_\ell, \bar{x}_\ell(\omega), \omega) \leq 0,\ \forall \omega \in \Omega\}
\end{aligned}
$$

Then, consider the following necessary assumptions about the functions involved in the problem (3).

A$_1$. Each of the objective functions F_i and $\bar{F}_i$ is assumed to be continuous and Caratheodory functions on $\mathbb{S}$, respectively (i.e., each F_i and $\bar{F}_i$ is continuous on $\mathbb{S}$ and $\bar{F}_i$ is measurable in ω) and convex with respect to their corresponding variables x_i and $\bar{x}_i$, where i indicates the ith level of decision hierarchy, $\forall i \in \{1, 2,\, \ldots,\, \ell\}$.

A$_2$. Each of the constraint functions G_i and $\bar{G}_i$ which determine the set $\mathbb{S}_i$, for $i \in \{1, 2,\, \ldots,\, \ell\}$, is assumed to be convex concerning their corresponding variables x_i, $\bar{x}_i$ and for all $\omega \in \Omega$; they are also Caratheodory functions.

Now, consider the bottom or ℓth level problem of (3), which is

$$\begin{aligned}
&\min_{x_\ell,\bar{x}_\ell} F_\ell(x_1,\, x_2,\, \ldots,\, x_\ell) + \mathbb{E}_\omega[\bar{F}_\ell(\bar{x}_1(\omega), \bar{x}_2(\omega),\, \ldots,\, \bar{x}_\ell(\omega), \omega)] \\
&\text{subject to } G_\ell(x_1,\, x_2,\, \ldots,\, x_\ell) \leq, \ldots, \\
&\qquad\qquad \bar{G}_\ell(x_1, \bar{x}_1(\omega), x_2, \bar{x}_2(\omega),\, \ldots,\, x_\ell, \bar{x}_\ell(\omega), \omega) \leq 0,\ \forall \omega \in \Omega
\end{aligned} \tag{4}$$

The involvement of the expectation function in the objective and the effect of the random variable ω in the constraint functions make problem (4) difficult to solve or make it even intractable. Therefore, it is customary to approximate the problem using the sample average approximation [38, 49]. The approximated form of problem (4) is given by

$$\begin{aligned}
&\min_{x_\ell,\bar{x}_\ell} F_\ell(x_1,\, x_2,\, \ldots,\, x_\ell) + \frac{1}{n}\sum_{i=1}^{n} \bar{F}_\ell(\bar{x}_1(\omega^i), \bar{x}_2(\omega^i),\, \ldots,\, \bar{x}_\ell(\omega^i), \omega^i) \\
&\text{subject to } (x_1, \bar{x}_1, x_2, \bar{x}_2,\, \ldots,\, x_\ell, \bar{x}_\ell) \in \mathbb{S}_\ell
\end{aligned} \tag{5}$$

where

$$\begin{aligned}
\mathbb{S}_\ell = \{(x_1, \bar{x}_1, x_2, \bar{x}_2,\, \ldots,\, x_\ell, \bar{x}_\ell) : {} & G_\ell(x_1,\, x_2,\, \ldots,\, x_\ell) \leq 0 \\
& \bar{G}_\ell(x_1, \bar{x}_1(\omega^i), x_2, \bar{x}_2(\omega^i),\, \ldots,\, x_\ell, \bar{x}_\ell(\omega^i), \omega^i) \leq 0,\ \forall \omega^i \in \Omega\}
\end{aligned} \tag{6}$$

and ω^i is a realization of the random variable having support $\hat{\Xi}$ using monte carlo simulation method.

If the variables in F_ℓ and $\bar{F}_\ell$ controlled by upper decision makers are considered as parameters (or known values) and if the objective functions F_ℓ and $\bar{F}_\ell$ are continuous and Caratheodory functions on $\mathbb{S}$, respectively, and convex concerning their corresponding variables x_ℓ and $\bar{x}_\ell$, and also if the constraint functions G_ℓ and $\bar{G}_\ell$ which determine the set $\mathbb{S}_\ell$ are convex concerning their corresponding variables $x_\ell, \bar{x}_\ell$ for all $\omega \in \Omega$ and they are also Caratheodory functions, then a problem solution (5) is an approximated solution

One can apply similar arguments on each of the hierarchical levels in the problem (3) to obtain an approximated ℓ-level deterministic problem

$$\begin{aligned}
&\min_{x_1,\bar{x}_1} \mathbb{F}_1(x_1, \bar{x}_1, x_2, \bar{x}_2,\, \ldots,\, x_\ell, \bar{x}_\ell) && \text{(7a)}\\
&\text{subject to } (x_1, \bar{x}_1, x_2, \bar{x}_2,\, \ldots,\, x_\ell, \bar{x}_\ell) \in \mathbb{S}_1, \\
&\quad \text{where } x_2, \bar{x}_2,\, \ldots,\, x_\ell \text{ and } \bar{x}_\ell \text{ solve} \\
&\quad \min_{x_2,\bar{x}_2} \mathbb{F}_2(x_1, \bar{x}_1, x_2, \bar{x}_2,\, \ldots,\, x_\ell, \bar{x}_\ell) \\
&\quad \text{subject to } (x_1, \bar{x}_1, x_2, \bar{x}_2, \ldots, x_\ell, \bar{x}_\ell) \in \mathbb{S}_2, \\
&\qquad \text{where } x_3, \bar{x}_3,\, \ldots,\, x_\ell \text{ and } \bar{x}_\ell \text{ solve} \\
&\qquad\qquad \ddots \\
&\qquad\qquad \text{where } x_\ell \text{ and } \bar{x}_\ell \text{ solve} \\
&\qquad\qquad \min_{x_\ell,\bar{x}_\ell} \mathbb{F}_\ell(x_1, \bar{x}_1, x_2, \bar{x}_2,\, \ldots,\, x_\ell, \bar{x}_\ell) && \text{(7b)}\\
&\qquad\qquad \text{subject to } (x_1, \bar{x}_1, x_2, \bar{x}_2,\, \ldots,\, x_\ell, \bar{x}_\ell) \in \mathbb{S}_\ell
\end{aligned}$$

where for each $j = 1,\, \ldots,\, \ell$,

$$\mathbb{F}_j(x_1, \bar{x}_1, x_2, \bar{x}_2, \ldots, x_\ell, \bar{x}_\ell) = F_j(x_1, x_2, \ldots, x_\ell) + \frac{1}{n}\sum_{i=1}^{n} \bar{F}_j(\bar{x}_1(\omega^i), \bar{x}_2(\omega^i), \ldots, \bar{x}_\ell(\omega^i), \omega^i) \tag{8}$$

and

$$\mathbb{S}_j = \{(x_1, \bar{x}_1, x_2, \bar{x}_2, \ldots, x_\ell, \bar{x}_\ell) : G_j(x_1, x_2, \ldots, x_\ell) \le 0 \\ \bar{G}_j(x_1, \bar{x}_1(\omega^i), x_2, \bar{x}_2(\omega^i), \ldots, x_\ell, \bar{x}_\ell(\omega^i), \omega^i) \le 0, \ \forall \omega^i \in \Omega\} \tag{9}$$

Since the problem solution (5) is the approximated solution for problem (4) [38, 49], the same argument can be generalized sequentially from the bottom level to the upper level. Therefore, the optimal solution of the approximated deterministic multilevel problem (i.e., problem (7)), which is constructed using a finite number of realizations of the random variable ω, is convergent to the optimal solution of TS-SMLP (problem (3)), as the number of realizations tends to infinity due to the law of large numbers, provided that assumptions $\mathbf{A}_1$ and $\mathbf{A}_2$ are satisfied.

Now, consider the wait-and-see variables as finite dimensional vectors $(\bar{x}_i^{(1)}, \bar{x}_i^{(2)}, \ldots, \bar{x}_i^{(n)}) = (\bar{x}_i(\omega^1), \bar{x}_i(\omega^2), \ldots, \bar{x}_i(\omega^n))$ such that $(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_\ell) \in \bar{X}(\neq \emptyset) \subseteq \mathbb{R}^{\bar{m}}$, where $\bar{x}_i^{(j)} \in [\bar{l}_i, \bar{u}_i]^{\bar{m}_i}$, $\bar{X}_i = [\bar{l}_i, \bar{u}_i]^{\bar{m}_i} \subseteq \mathbb{R}^{\bar{m}_i}$, $l_i^j(\bar{l}_i^j)$ and $u_i^j(\bar{u}_i^j)$ are lower and upper bounds for each decision component of the hear-and-now variables and the wait-and-see variables, respectively, $\sum_{i=1}^{\ell}(m_i + \bar{m}_i) = m$ and $\mathbb{K}^i := \prod_{t=i}^{\ell}(X_t \times \bar{X}_t)$. Below some important TS-SMLP definitions are presented.

Definition 1. For the ℓ level TS-SMLP problem (7)

1. The relaxed constraint region is given by

$$\Psi = \{(x_1, \bar{x}_1, \ldots, x_\ell, \bar{x}_\ell) \in \mathbb{S} : (x_1, \bar{x}_1, \ldots, x_\ell, \bar{x}_\ell) \in \mathbb{S}_1 \cap \mathbb{S}_2 \cap \cdots \cap \mathbb{S}_\ell\}$$

2. For each given vector $(x_1^\star, \bar{x}_1^\star, \ldots, x_{i-1}^\star, \bar{x}_{i-1}^\star) \in \prod_{t=1}^{i-1}(X_t \times \bar{X}_t)$, $2 \le i \le \ell$, the ith level feasible region is given by

$$\Psi(x_1^\star, \bar{x}_1^\star, \ldots, x_{i-1}^\star, \bar{x}_{i-1}^\star) = \{(x_i, \bar{x}_i, \ldots, x_\ell, \bar{x}_\ell) \in \prod_{t=i}^{\ell}(X_t \times \bar{X}_t) : \\ (x_1^\star, \bar{x}_1^\star, \ldots, x_{i-1}^\star, \bar{x}_{i-1}^\star, x_i, \bar{x}_i, \ldots, x_\ell, \bar{x}_\ell) \in \mathbb{S}_i \cap \cdots \cap \mathbb{S}_\ell\}$$

3. Projection of Ψ on to the decision space of the 1st, ..., ith, $1 \le i < \ell$, level is given by

$$\Psi^i = \{(x_1, \bar{x}_1, \ldots, x_i, \bar{x}_i) \in \prod_{t=1}^{i}(X_t \times \bar{X}_t) : \\ \exists (x_{i+1}, \bar{x}_{i+1}, \ldots, x_\ell, \bar{x}_\ell) \in \prod_{t=i+1}^{\ell}(X_t \times \bar{X}_t) \text{ with } (x_1, \bar{x}_1, \ldots, x_\ell, \bar{x}_\ell) \in \Psi\}$$

Note that $\Psi^\ell = \Psi$.

4. For each $(x_1^\star, \bar{x}_1^\star, \dots, x_{\ell-1}^\star, \bar{x}_{\ell-1}^\star) \in \Psi^{\ell-1}$, the rational reaction set for the ℓth level is given by (which is modified for stochastic case from literature [59])

$$\begin{aligned}\Phi^\ell(x_1^\star, \bar{x}_1^\star, \dots, x_{\ell-1}^\star, \bar{x}_{\ell-1}^\star) = \operatorname*{argmin}_{x_\ell, \bar{x}_\ell} \{ &\mathbb{F}_\ell(x_1^\star, \bar{x}_1^\star, \dots, x_{\ell-1}^\star, \bar{x}_{\ell-1}^\star, x_\ell, \bar{x}_\ell) : \\ &(x_\ell, \bar{x}_\ell) \in \mathbb{K}^\ell, (x_1^\star, \bar{x}_1^\star, \dots, x_{\ell-1}^\star, \bar{x}_{\ell-1}^\star, x_\ell, \bar{x}_\ell) \in \Psi \}\end{aligned}$$

Then inductively, for $2 \le i \le \ell - 1$ and $(x_1^\star, \bar{x}_1^\star, \dots, x_{i-1}^\star, \bar{x}_{i-1}^\star) \in \Psi^{i-1}$, define the rational reaction set to be

$$\begin{aligned}\Phi^i(x_1^\star, \bar{x}_1^\star, \dots, x_{i-1}^\star, \bar{x}_{i-1}^\star) = \operatorname*{argmin}_{x_i, \bar{x}_i} \{ &\mathbb{F}_i(x_1^\star, \bar{x}_1^\star, \dots, x_{i-1}^\star, \bar{x}_{i-1}^\star, x_i, \bar{x}_i, \dots, x_\ell, \bar{x}_\ell) : \\ &(x_1^\star, \bar{x}_1^\star, \dots, x_{i-1}^\star, \bar{x}_{i-1}^\star, x_i^\circ, \bar{x}_i^\circ, \dots, x_\ell^\circ, \bar{x}_\ell^\circ) \in \Psi, \\ &(x_t, \bar{x}_t, \dots, x_\ell, \bar{x}_\ell) \in \Phi^t(x_1^\star, \bar{x}_1^\star, \dots, x_{i-1}^\star, \bar{x}_{i-1}^\star, x_i, \bar{x}_i, \dots, x_{t-1}, \bar{x}_{t-1}), \\ &\forall i + 1 \le t \le \ell - 1 \text{ and } (x_i, \bar{x}_i, \dots, x_\ell, \bar{x}_\ell) \in \mathbb{K}^i \}\end{aligned} \tag{10}$$

5. The induced region for the leader's decision is defined as

$$\begin{aligned}\hat{\Upsilon} = \{ &(x_1, \bar{x}_1, \dots, x_\ell, \bar{x}_\ell) \in \mathbb{S}_1 : (x_1, \bar{x}_1) \in X_1 \times \bar{X}_1 \\ &(x_2, \bar{x}_2, \dots, x_\ell, \bar{x}_\ell) \in \Phi^2(x_1, \bar{x}_1) \}\end{aligned}$$

The TS-SMLP problem (7) can be equivalently [20] described as a single level problem using the set $\hat{\Upsilon}$ (if it is nonempty) as

$$\min_{(x_1, \bar{x}_1, \dots, x_\ell, \bar{x}_\ell) \in \hat{\Upsilon}} \mathbb{F}_1(x_1, \bar{x}_1, x_2, \bar{x}_2, \dots, x_\ell, \bar{x}_\ell) \tag{11}$$

Therefore, any optimal solution of problem (11) is a Stackelberg solution for TS-SMLP given in problem (7) because of this equivalence.

In problem (11), the leader cannot decide on all of the components of the decision vector $(x_1, \bar{x}_1, \dots, x_\ell, \bar{x}_\ell)$ but only on $(x_1, \bar{x}_1)$, which is the decision sub-vector for the leader's objective function, $\mathbb{F}_1(x_1, \bar{x}_1, \dots, x_\ell, \bar{x}_\ell)$ and similarly, $(x_i, \bar{x}_i)$ is the decision sub-vector for the ith hierarchical level decision maker, where $1 \le i \le \ell$. Assume that actions or decisions are made sequentially beginning from the top-level decision maker who has control over the sub-vector $(x_1, \bar{x}_1) \in X_1 \times \bar{X}_1$, followed by the 2nd level decision maker who has control over the sub-vector $(x_2, \bar{x}_2) \in X_2 \times \bar{X}_2$ down through the ℓth level decision maker who has control over the sub-vector $(x_\ell, \bar{x}_\ell) \in X_\ell \times \bar{X}_\ell$. If the ℓth level decision maker has more than one minimizer, then the decision maker at $(\ell - 1)$th level needs a clear definition of whether he/she uses an optimistic or pessimistic approach and the same approach is used for the remaining hierarchical levels of decision.

On the other hand, corresponding to each choice of a decision variable from above, the reaction is also made sequentially beginning from the bottom level decision maker up through the 2nd level decision maker, this procedure continues sequentially until they attain the Stackelberg equilibrium.

In the case when the second-level and/or subsequent lower-level subproblems of (7) have multiple optimal solutions, we assume that the decision maker at the respective level chooses the one that leads to the best value for the upper-level decision maker, i.e., we consider the sequential optimistic version of multilevel optimization problems.

Now consider the following additional necessary assumptions to ensure the existence of an optimal problem solution (7).

- $\mathbf{A}_3$. Assumption $\mathbf{A}_1$ is employed after all the random variables are realized and problem (3) is approximated using sample average approximation.
- $\mathbf{A}_4$. Assumption $\mathbf{A}_2$ is employed after all the random variables are realized and problem (3) is approximated using sample average approximation.
- $\mathbf{A}_5$. All the decision variables are assumed to be in a closed boxed region.
- $\mathbf{A}_6$. The optimal reaction set in problem (7) at any hierarchical level is assumed to be non-empty.

Under the consideration of the above assumptions, the following theorem shows the existence of a problem solution (7).

Theorem 1. If assumptions $\mathbf{A}_3$–$\mathbf{A}_6$ hold, then problem (7) has an optimal solution.

Proof. Decision is assumed to be started from the top level and sequentially cascades down through the bottom level decision maker. Corresponding to any choice $(x_1, \bar{x}_1, \ldots, x_{\ell-1}, \bar{x}_{\ell-1})$ of the top $(\ell - 1)$ decision makers, there is at least one reaction $(x_\ell, \bar{x}_\ell)$ by $\mathbf{A}_6$. Here, one of the elements can be possibly selected in a unique way from the rational reaction set $\Phi^\ell(x_1, \bar{x}_1, \ldots, x_{\ell-1}, \bar{x}_{\ell-1})$ due to the sequential optimistic assumption of the problem and say the value is $\Phi^{\ell\circ}(x_1, \bar{x}_1, \ldots, x_{\ell-1}, \bar{x}_{\ell-1})$. In a similar argument, the optimal reaction can be selected at each hierarchical level in a unique way from the bottom level up through the top level. Let $\Phi^{2\circ}(x_1, \bar{x}_1) := (x_2, \bar{x}_2, \ldots, x_\ell, \bar{x}_\ell)$ be the chosen unique optimal reaction for the 1st level decision maker for his/her decision $(x_1, \bar{x}_1)$ and then in terms of this value, problem (11) becomes

$$\min_{\substack{x_1, \bar{x}_1 \\ (x_1, \bar{x}_1, \ldots, x_\ell, \bar{x}_\ell) \in \hat{\Upsilon}}} \quad \mathbb{F}_1(x_1, \bar{x}_1, \Phi^{2\circ}(x_1, \bar{x}_1)) \tag{12}$$

Therefore, since $\hat{\Upsilon}$ is closed because of assumption $\mathbf{A}_4$ and $\mathbf{A}_5$, and because of $\hat{\Upsilon} \subseteq \mathbb{S}$ and $\mathbb{S}$ is a bounded set, the set $\hat{\Upsilon}$ is compact. Hence, the Weierstrass theorem implies that there exists a solution for problem (12) due to the continuity assumption in $\mathbf{A}_3$ over the compact set $\hat{\Upsilon}$. □

3. Solution approach

3.1. Two-stage stochastic programming problems

The SSE solution procedure is a non-derivative meta-heuristic type approach. The solution procedure was proposed for solving SBLP problems in [25] for a single follower case and is extended in [26] for multiple followers problems. In this work, it is further extended for solving SMLP problems for any number of hierarchical levels. The approach is based on the realization of the random variables and then selecting an optimal reaction from the partitioned region using the systematic sampling approach at each hierarchical level sequentially from the bottom level up through the top level. We implement the PSO algorithm for searching for a better optimal reaction if there is one at each hierarchical level. Parameters that are used in the approach are described in Table 1.

Table 1. Description of parameters used in the algorithm

Parameter	Description
d_i^j	length of each side of the hyper-box (hypercube) along the jth component
β	parameter of the partition
a_i^j	$a_i^j = \dfrac{d_i^j}{\beta}, \forall j \in \{1, \ldots, m_i\}, \forall i \in \{1, \ldots, \ell\}$
b_i^j	two consecutive a_i^j partition are merged to be b_i^j as $b_i^j(p) = a_i^j(2p-1) + a_i^j(2p)$, where total number of merged partitions is represented by p
ρ	SSE algorithm maximum iteration number
β°	a parameter $\beta^\circ = \dfrac{\beta}{\rho}$ is used to avoid repetition of representatives at each iteration
q	particles number
π	PSO maximum iteration number
$l_i^j, \bar{l}_i^j$	lower bounds for each decision component of "hear and now variables" and "wait and see variables"
$u_i^j, \bar{u}_i^j$	upper bounds for each decision component of "hear and now variables" and "wait and see variables"

After realizing the random variables, the next step is partitioning the leader's decision spaces using a systematic sampling approach as presented in Figure 1. The schematic diagram indicates how the hierarchical decision spaces at each level can be partitioned using a systematic sampling approach. A box in each of the $\mathbf{F}_i$'s represents one hyper-box[1] (or hypercube) in the respective dimension. One representative is randomly chosen from each of the partitioned regions and then the rest representatives of the actions are selected systematically as described in reference [25]. Using a similar mechanism, each of the remaining hierarchical decision spaces (DS) as presented in Figure 1 can be partitioned using a systematic sampling approach and two strategies $x_i^j(1) = l_i^j + \text{rand}(1)\beta$ and $x_i^j(2) = l_i^j + \text{rand}(1)\beta + \beta$, respectively, are selected from the first two consecutive partitions. The other representatives of the decision variables from the remaining partitions are selected similarly as outlined in [25] for each hierarchical level i.

In a hierarchical decision system, a decision process starts from the top-level decision maker by anticipating the potential reaction of the followers sequentially. The intermediate decision maker $\mathbb{F}_i$, $2 \leq i \leq \ell - 1$, optimizes his/her problem in light of the knowledge of the upper $(i-1)$ decision makers' decision and the potential reactions of the bottom decision makers. Lastly, given the top $(\ell - 1)$ decision makers' decision, the ℓth level decision maker decides to optimize his/her objective function. And response of an action or a strategy from each partition region at each hierarchical decision space starts from the bottom level up through the top level. This decision process is repeatedly carried out until Stackelberg equilibrium is attained in all vertical structures at the current iteration.

[1]By hyper-box (or hypercube) we refer to a geometric structure in an arbitrary finite-dimensional real space, which represents the n-dimensional analog of a square in $\mathbb{R}^2$, and a cube in $\mathbb{R}^3$.

Figure 1. Partitioning each hierarchical level decision space

However, there might be another better solution for the problem because of the possible infinite representatives from each decision space. So, one keeps this solution and selects another set of possible combinations of solutions from each sub-region. But the repetition of representative actions has to be avoided from any of the sub-regions of the leader's decision space by further subdividing each partition region into a maximum number of iteration ρ as described in reference [25]. The procedure of SSE for solving the TS-SMLP problem is presented using a flowchart in Figure 2.

The pseudo-code is given in Algorithm 1 for solving the reformulated TS-SMLP as a deterministic multilevel program**Algorithm 1**. Solution procedure pseudo code

Step 1. A multilevel programming problem with bounded decision variables is considered.

Input: $m_i, \bar{m}_i, l_i^j, u_i^j, \bar{l}_i^{\bar{j}}, \bar{u}_i^{\bar{j}}, \mathbb{S}_i, \mathbb{F}_i, d_i^j, \bar{d}_i^{\bar{j}}, \forall j \in \{1, \ldots, m_i\}, \forall \bar{j} \in \{1, \ldots, \bar{m}_i\}, \forall i \in \{1, \ldots, \ell\}$.

Algorithm parameters: $\rho, \beta, \beta^\circ = \frac{\beta}{\rho}, \breve{a}_i^j = 2 \cdot \text{ceil}(\frac{d_i^j}{2\beta}), \breve{b}_i^j = \text{floor}(\frac{\breve{a}_i^j}{2})$ and $\Gamma = \text{randperm}(\rho)$, where $\breve{b}_i^j$ is the number of partitions of b_i^j and $\breve{a}_i^j$ is the number of partitions of a_i^j. For wait-and-see variables is also done the same.

Step 2. Iteration counter is set to be $\kappa = 1$, and initialize the set of solutions $\Upsilon = \emptyset$.

Step 3. Leader's decision space is partitioned, given by the hyper-box $[l_1, u_1]^{m_1}$.

$x_1^j(1) = l_1^j + (\Gamma(\kappa) - 1)\beta^\circ + \text{rand}(1)\beta^\circ$
$x_1^j(2) = l_1^j + \beta + (\rho - \Gamma(\kappa) + 1)\beta^\circ + \text{rand}(1)\beta^\circ$
for $q = 3$ to $2\breve{b}_1 - 1$ using 2 steps do
  $x_1^j(q) = x_1^j(q-2) + 2\beta$
  $x_1^j(q+1) = x_1^j(q-1) + 2\beta$
end for

The same is done for the wait-and-see-variables space (which is also assumed to be a hyper-box). In addition, partition each hierarchical decision space on the hyper-box $[l_i, u_i]^{m_i}$.

for $i = 2$ to ℓ
  $x_i^j(1) = l_i^j + \text{rand}(1)\beta$
  $x_i^j(2) = l_i^j + \text{rand}(1)\beta + \beta$
  for $q = 3$ to $2\breve{b}_1 - 1$ using 2 steps do
    $x_i^j(q) = x_2^j(q-2) + 2\beta$
    $x_i^j(q+1) = x_i^j(q-1) + 2\beta$
  end for q
end for i

The same procedure is applied also for the wait-and-see variables of the follower's problem.

Step 4. For $i = 1, \ldots, \ell - 1$, and for each $(x_1^\star, \bar{x}_1^\star, \ldots, x_i^\star, \bar{x}_i^\star) \in X_1 \times \bar{X}_1 \times \cdots \times X_i \times \bar{X}_i$, find an optimal reaction sequentially from the bottom level to 2nd level using Definition 1.4. For this chosen vector,

- if an optimal reaction is obtained, go to Step 5.
- if no optimal reaction is obtained, set $\kappa := \kappa + 1$ and go to Step 3.

Step 5. If $\kappa = 1$, then set $\Upsilon := \Upsilon \cup \{(x_1^\kappa, \bar{x}_1^\kappa, x_2^\kappa, \bar{x}_2^\kappa, \ldots, x_\ell^\kappa, \bar{x}_\ell^\kappa)\}$. Otherwise (i.e., if $\kappa \geq 2$), select a best solution using

$$\mathbb{F}_1(x_1^\kappa, \bar{x}_1^\kappa, x_2^\kappa, \bar{x}_2^\kappa, \ldots, x_\ell^\kappa, \bar{x}_\ell^\kappa) \leq \mathbb{F}_1(x_1, \bar{x}_1, x_2, \bar{x}_2, \ldots, x_\ell, \bar{x}_\ell) \tag{13}$$

such that $(x_1, \bar{x}_1, x_2, \bar{x}_2, \ldots, x_\ell, \bar{x}_\ell) \in \Upsilon$ and then set $(x_1^\kappa, \bar{x}_1^\kappa, x_2^\kappa, \bar{x}_2^\kappa, \ldots, x_\ell^\kappa, \bar{x}_\ell^\kappa)$ as the best solution for iteration κ and update $\Upsilon := \Upsilon \cup \{(x_1^\kappa, \bar{x}_1^\kappa, x_2^\kappa, \bar{x}_2^\kappa, \ldots, x_\ell^\kappa, \bar{x}_\ell^\kappa)\}$. Terminate, if $\rho = \kappa$, else set $\kappa := \kappa + 1$ and go to Step 3.

Step 6. Output: $(x_1^\rho, \bar{x}_1^\rho, x_2^\rho, \bar{x}_2^\rho, \ldots, x_\ell^\rho, \bar{x}_\ell^\rho) := (x_1^\circ, \bar{x}_1^\circ, x_2^\circ, \bar{x}_2^\circ, \ldots, x_\ell^\circ, \bar{x}_\ell^\circ)$.

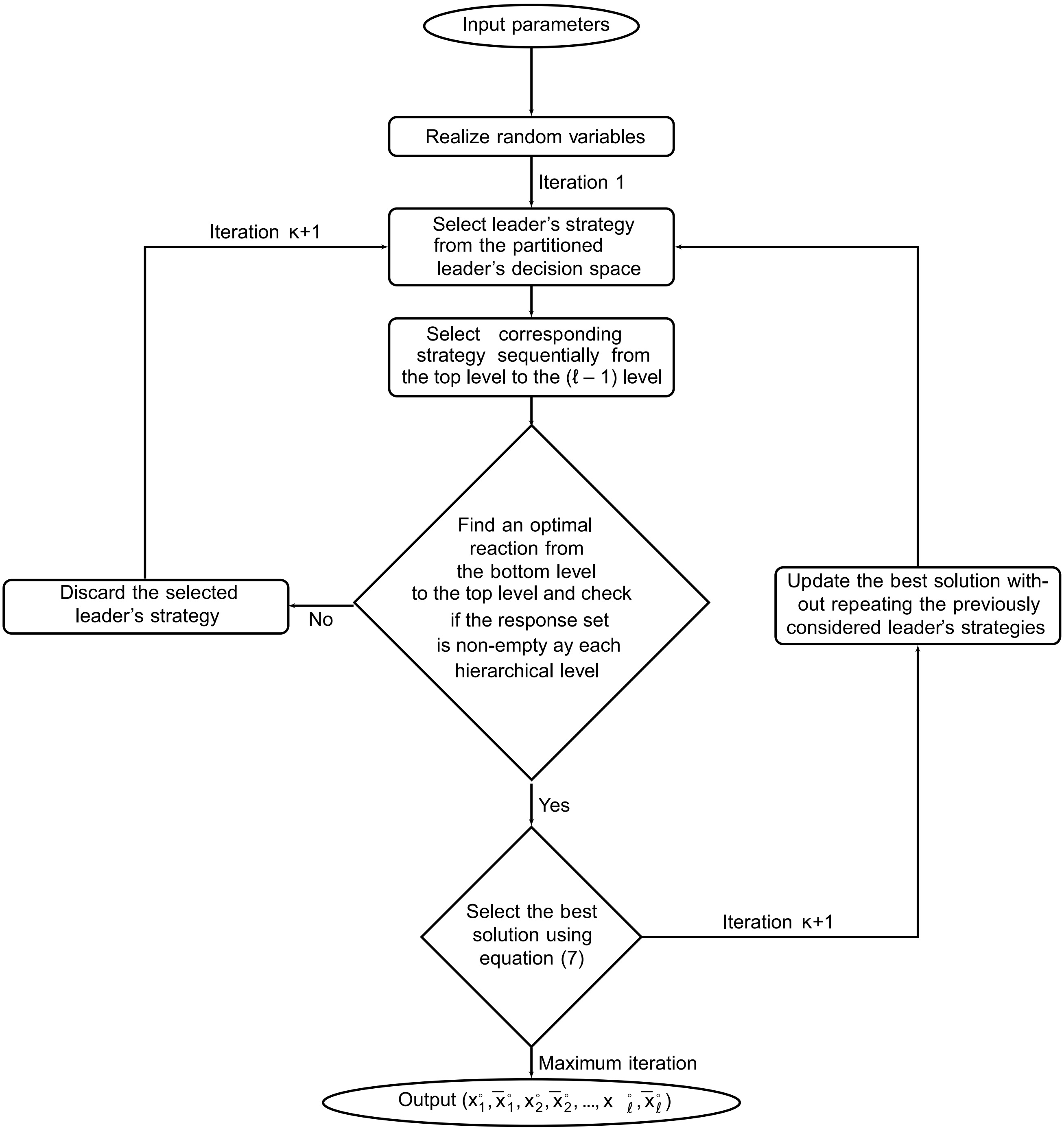

Figure 2. Flowchart for solving TS-SMLP

Now, the convergence of the above iterative procedures which is described by Algorithm 1 is justified in the following theorem.

Theorem 2. For problem (7), let assumptions $\mathbf{A}_3$–$\mathbf{A}_6$ hold and let $(x_1^\circ,\ \bar{x}_1^\circ, x_2^\circ, \bar{x}_2^\circ,\ \ldots,\ x_\ell^\circ,\ \bar{x}_\ell^\circ)$ be a Stackelberg solution for problem (7). Then, for any chosen parameter $\beta > 0$ the iterations of Algorithm 1 produce a sequence of points $(x_1^\kappa, \bar{x}_1^\kappa, x_2^\kappa, \bar{x}_2^\kappa,\ \ldots,\ x_\ell^\kappa, \bar{x}_\ell^\kappa)$ that converges to $(x_1^\circ, \bar{x}_1^\circ, x_2^\circ, \bar{x}_2^\circ, \ldots, x_\ell^\circ, \bar{x}_\ell^\circ)$.

Proof. In the case of $\ell = 2$, the convergence of the Stackelberg equilibrium was shown in [25]. Extending the result in [25] for any hierarchical level ℓ and using similar argument, $(x_1^\kappa, \bar{x}_1^\kappa, x_2^\kappa, \bar{x}_2^\kappa,\ \ldots,\ x_\ell^\kappa, \bar{x}_\ell^\kappa) \to (x_1^\circ, \bar{x}_1^\circ, x_2^\circ, \bar{x}_2^\circ,\ \ldots,\ x_\ell^\circ, \bar{x}_\ell^\circ)$ is wanted to be shown as $\kappa \to \infty$. Note that each of $(x_1^\kappa, \bar{x}_1^\kappa, x_2^\kappa, \bar{x}_2^\kappa,\ \ldots,\ x_\ell^\kappa, \bar{x}_\ell^\kappa)$ and $(x_1^\circ, \bar{x}_1^\circ, x_2^\circ, \bar{x}_2^\circ,\ \ldots,\ x_\ell^\circ, \bar{x}_\ell^\circ)$ satisfy Definition 1. optimal reaction condition (4) at any hierarchical level sequentially starting from the last level up through the first level. The Euclidean distance between $(x_1^\kappa, \bar{x}_1^\kappa, x_2^\kappa, \bar{x}_2^\kappa,\ \ldots,\ x_\ell^\kappa, \bar{x}_\ell^\kappa)$ and $(x_1^\circ, \bar{x}_1^\circ, x_2^\circ, \bar{x}_2^\circ,\ \ldots,\ x_\ell^\circ, \bar{x}_\ell^\circ)$ is given by $\|(x_1^\kappa, \bar{x}_1^\kappa, x_2^\kappa, \bar{x}_2^\kappa,\ \ldots,\ x_\ell^\kappa, \bar{x}_\ell^\kappa) - (x_1^\circ, \bar{x}_1^\circ, x_2^\circ, \bar{x}_2^\circ,\ \ldots,\ x_\ell^\circ, \bar{x}_\ell^\circ)\|$ such that

$$
\begin{aligned}
&\|(x_1^\kappa, \bar{x}_1^\kappa, x_2^\kappa, \bar{x}_2^\kappa,\ \ldots,\ x_\ell^\kappa, \bar{x}_\ell^\kappa) - (x_1^\circ, \bar{x}_1^\circ, x_2^\circ, \bar{x}_2^\circ,\ \ldots,\ x_\ell^\circ, \bar{x}_\ell^\circ)\|^2 && \text{(14a)}\\
&\quad \le \|(x_1^\kappa, \bar{x}_1^\kappa) - (x_1^\circ, \bar{x}_1^\circ)\|^2 + \|(x_2^\kappa, \bar{x}_2^\kappa) - (x_2^\circ, \bar{x}_2^\circ)\|^2 + \cdots + \|(x_\ell^\kappa, \bar{x}_\ell^\kappa) - (x_\ell^\circ, \bar{x}_\ell^\circ)\|^2\\
&\quad = \|(x_1^\kappa, \bar{x}_1^\kappa) - (x_1^\circ, \bar{x}_1^\circ)\|^2\\
&\qquad + \|\operatorname*{argmin}_{x_2, \bar{x}_2} \mathbb{F}_2(x_1^\kappa, \bar{x}_1^\kappa, x_2(x_1^\kappa, \bar{x}_1^\kappa), \bar{x}_2(x_1^\kappa, \bar{x}_1^\kappa),\ \ldots,\ x_\ell(x_1^\kappa, \bar{x}_1^\kappa), \bar{x}_\ell(x_1^\kappa, \bar{x}_1^\kappa))\\
&\qquad - \operatorname*{argmin}_{x_2, \bar{x}_2} \mathbb{F}_2(x_1^\circ, \bar{x}_1^\circ, x_2(x_1^\circ, \bar{x}_1^\circ), \bar{x}_2(x_1^\circ, \bar{x}_1^\circ),\ \ldots,\ x_\ell(x_1^\circ, \bar{x}_1^\circ), \bar{x}_\ell(x_1^\circ, \bar{x}_1^\circ))\|^2\\
&\qquad + \cdots + \|\operatorname*{argmin}_{x_{\ell-1}, \bar{x}_{\ell-1}} \mathbb{F}_{\ell-1}(x_1^\kappa, \bar{x}_1^\kappa,\ \ldots,\ x_{\ell-2}^\kappa, \bar{x}_{\ell-2}^\kappa, x_{\ell-1}(x_1^\kappa, \bar{x}_1^\kappa), \bar{x}_{\ell-1}(x_1^\kappa, \bar{x}_1^\kappa), x_\ell(x_1^\kappa, \bar{x}_1^\kappa), \bar{x}_\ell(x_1^\kappa, \bar{x}_1^\kappa))\\
&\qquad - \operatorname*{argmin}_{x_{\ell-1}, \bar{x}_{\ell-1}} \mathbb{F}_{\ell-1}(x_1^\circ, \bar{x}_1^\circ,\ \ldots,\ x_{\ell-2}^\circ, \bar{x}_{\ell-2}^\circ, x_{\ell-1}(x_1^\circ, \bar{x}_1^\circ), \bar{x}_{\ell-1}(x_1^\circ, \bar{x}_1^\circ), x_\ell(x_1^\circ, \bar{x}_1^\circ)), \bar{x}_\ell(x_1^\circ, \bar{x}_1^\circ))\|^2\\
&\qquad + \|\operatorname*{argmin}_{x_\ell, \bar{x}_\ell} \mathbb{F}_\ell(x_1^\kappa, \bar{x}_1^\kappa,\ \ldots,\ x_{\ell-1}^\kappa, \bar{x}_{\ell-1}^\kappa, x_\ell(x_1^\kappa, \bar{x}_1^\kappa), \bar{x}_\ell(x_1^\kappa, \bar{x}_1^\kappa))\\
&\qquad - \operatorname*{argmin}_{x_\ell, \bar{x}_\ell} \mathbb{F}_\ell(x_1^\circ, \bar{x}_1^\circ,\ \ldots,\ x_{\ell-1}^\circ, \bar{x}_{\ell-1}^\circ, x_\ell(x_1^\circ, \bar{x}_1^\circ), \bar{x}_\ell(x_1^\circ, \bar{x}_1^\circ))\|^2\\
&\quad \le \|x_1^\kappa - x_1^\circ\|^2 + \|\bar{x}_1^\kappa - \bar{x}_1^\circ\|^2 + && \text{(14b)}\\
&\qquad \|\operatorname*{argmin}_{x_2, \bar{x}_2} \mathbb{F}_2(x_1^\kappa, \bar{x}_1^\kappa, x_2(x_1^\kappa, \bar{x}_1^\kappa), \bar{x}_2(x_1^\kappa, \bar{x}_1^\kappa),\ \ldots,\ x_\ell(x_1^\kappa, \bar{x}_1^\kappa), \bar{x}_\ell(x_1^\kappa, \bar{x}_1^\kappa))\\
&\qquad - \operatorname*{argmin}_{x_2, \bar{x}_2} \mathbb{F}_2(x_1^\circ, \bar{x}_1^\circ, x_2(x_1^\circ, \bar{x}_1^\circ), \bar{x}_2(x_1^\circ, \bar{x}_1^\circ),\ \ldots,\ x_\ell(x_1^\circ, \bar{x}_1^\circ), \bar{x}_\ell(x_1^\circ, \bar{x}_1^\circ))\|^2\\
&\qquad + \cdots + \|\operatorname*{argmin}_{x_{\ell-1}, \bar{x}_{\ell-1}} \mathbb{F}_{\ell-1}(x_1^\kappa, \bar{x}_1^\kappa,\ \ldots,\ x_{\ell-2}^\kappa, \bar{x}_{\ell-2}^\kappa, x_{\ell-1}(x_1^\kappa, \bar{x}_1^\kappa), \bar{x}_{\ell-1}(x_1^\kappa, \bar{x}_1^\kappa), x_\ell(x_1^\kappa, \bar{x}_1^\kappa), \bar{x}_\ell(x_1^\kappa, \bar{x}_1^\kappa))\\
&\qquad - \operatorname*{argmin}_{x_{\ell-1}, \bar{x}_{\ell-1}} \mathbb{F}_{\ell-1}(x_1^\circ, \bar{x}_1^\circ,\ \ldots,\ x_{\ell-2}^\circ, \bar{x}_{\ell-2}^\circ, x_{\ell-1}(x_1^\circ, \bar{x}_1^\circ), \bar{x}_{\ell-1}(x_1^\circ, \bar{x}_1^\circ), x_\ell(x_1^\circ, \bar{x}_1^\circ)), \bar{x}_\ell(x_1^\circ, \bar{x}_1^\circ))\|^2\\
&\qquad + \|\operatorname*{argmin}_{x_\ell, \bar{x}_\ell} \mathbb{F}_\ell(x_1^\kappa, \bar{x}_1^\kappa,\ \ldots,\ x_{\ell-1}^\kappa, \bar{x}_{\ell-1}^\kappa, x_\ell(x_1^\kappa, \bar{x}_1^\kappa), \bar{x}_\ell(x_1^\kappa, \bar{x}_1^\kappa))\\
&\qquad - \operatorname*{argmin}_{x_\ell, \bar{x}_\ell} \mathbb{F}_\ell(x_1^\circ, \bar{x}_1^\circ,\ \ldots,\ x_{\ell-1}^\circ, \bar{x}_{\ell-1}^\circ, x_\ell(x_1^\circ, \bar{x}_1^\circ), \bar{x}_\ell(x_1^\circ, \bar{x}_1^\circ))\|^2
\end{aligned}
$$

In the inequality (14b) the norm of the here-and-now variables of the leader is given by

$$\|x_1^\kappa - x_1^\circ\|^2 = \sum_{j=1}^{m_1} |x_1^{j(\kappa)} - x_1^{j(\circ)}|^2 \tag{15}$$

Then, in equation (15), the Stackelberg solution component corresponding to the chosen parameter β, $x_1^{j(\circ)}$ is assumed to be in the middle of the partition. Using Theorem 2 in [25], $x_1^{j(\kappa)} \to x_1^{j(\circ)}$ as $\kappa \to \infty$

and/or for $\beta \to 0$. This argument is also true if $x_1^{j(\circ)}$ is at any position of the partition since only their corresponding distance is considered. So, $x_1^\kappa \to x_1^\circ$. One can show using similar arguments that $\bar{x}_i^\kappa \to \bar{x}_i^\circ$ for all $i = 1, \ldots, \ell$ in the inequality (14b) since each strategy is a sequential action and reaction of the other decision maker's strategy.

Therefore, $(x_1^\kappa, \bar{x}_1^\kappa, x_2^\kappa, \bar{x}_2^\kappa, \ldots, x_\ell^\kappa, \bar{x}_\ell^\kappa) \to (x_1^\circ, \bar{x}_1^\circ, x_2^\circ, \bar{x}_2^\circ, \ldots, x_\ell^\circ, \bar{x}_\ell^\circ)$ as $\kappa \to \infty$. $\square$

From Theorem 2, $(x_1^\kappa, \bar{x}_1^\kappa, x_2^\kappa, \bar{x}_2^\kappa, \ldots, x_\ell^\kappa, \bar{x}_\ell^\kappa)$ converges to $(x_1^\circ, \bar{x}_1^\circ, x_2^\circ, \bar{x}_2^\circ, \ldots, x_\ell^\circ, \bar{x}_\ell^\circ)$ as $\kappa \to \infty$ means that each hierarchical decision maker's functional value at each level is convergent due to the continuity assumption in $\mathbf{A}_3$. The convergence proof for the SSE solution procedure is done only for problem (7) fulfilling the assumptions raised in Theorem 2.

However, there are an infinite number of strategies between two consecutive systematically selected strategies at each hierarchical level's decision space which is shown in Figure 1. So, there might be a better solution or strategy out of those infinite strategies due to the continuity property at each hierarchical decision space. In addition to this, the bottom-level problem is a standard mathematical programming problem for each strategy of the top $(\ell - 1)$ decision-makers decision. For implementation purposes, the PSO technique is used for solving the bottom-level decision maker's problem and searching for a better optimal reaction, if it exists, at each hierarchical level's decision space. The better solution from the sampled partition region is used as an initialization for PSO.

The penalty method is used to control infeasible solutions by converting the problem at each hierarchical level into an unconstrained problem by adding a penalty function to the objective function. The penalty function consists of a penalty parameter Λ multiplied by a measure of violation of the constraints K. The measure of violation is non-zero when the constraint is violated and is zero in the region where the constraint is not violated. The converted unconstrained problem is given in the following equation for each hierarchy level i, where $i \in \{1, 2, \ldots, \ell\}$.

$$\hat{\mathbb{F}}_i(x_1, \bar{x}_1, x_2, \bar{x}_2, \ldots, x_\ell, \bar{x}_\ell) = \mathbb{F}_i(x_1, \bar{x}_1, x_2, \bar{x}_2, \ldots, x_\ell, \bar{x}_\ell) + \Lambda \cdot \max\{K, 0\} \tag{16}$$

where $K = 0$ if $(x_1, \bar{x}_1, x_2, \bar{x}_2, \ldots, x_\ell, \bar{x}_\ell) \in \mathbb{S}_i$ and $K =$ some large constant number if $(x_1, \bar{x}_1, x_2, \bar{x}_2, \ldots, x_\ell, \bar{x}_\ell) \notin \mathbb{S}_i$.

3.2. Chance constrained stochastic programming problems

The solution procedure described in the Subsection 3.1 also works for other types of problems like CC-MLP, where the chance constraint programming problem is defined at each hierarchical level of the multilevel programming problem. Mathematical formulation for CC-MLP problem $\forall \omega \in \Omega$ is given by

$$\min_{\bar{x}_1} \gamma_1 \mathbb{E}_\omega[\bar{F}_1(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_\ell, \omega)] + \gamma_2 \mathrm{Var}_\omega[\bar{F}_1(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_\ell, \omega)] \tag{17a}$$

$$\text{subject to } P\{\bar{G}_1(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_\ell, \omega) \leq 0\} \geq \alpha,$$

$$\text{where } \bar{x}_2, \bar{x}_3, \ldots, \bar{x}_{\ell-1} \text{ and } \bar{x}_\ell \text{ solve}$$

$$\min_{\bar{x}_2} \gamma_1 \mathbb{E}_\omega[\bar{F}_2(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_\ell, \omega)] + \gamma_2 \mathrm{Var}_\omega[\bar{F}_2(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_\ell, \omega)]$$

$$\text{subject to } \mathrm{P}\{_2(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_\ell, \omega) \leq 0\} \geq \alpha,$$
$$\text{where } \bar{x}_3, \bar{x}_4, \ldots, \bar{x}_{\ell-1} \text{ and } \bar{x}_\ell \text{ solve}$$
$$\ddots$$
$$\min_{\bar{x}_i} \gamma_1 \mathbb{E}_\omega[\bar{F}_i(\bar{x}_1, \ldots, \bar{x}_\ell, \omega)] + \gamma_2 \mathrm{Var}_\omega[\bar{F}_i(\bar{x}_1, \ldots, \bar{x}_\ell, \omega)] \quad (17a)$$
$$\text{subject to } \mathrm{P}\{_i(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_\ell, \omega) \leq 0\} \geq \alpha,$$
$$\text{where } \bar{x}_{i+1}, \bar{x}_{i+2}, \ldots, \bar{x}_{\ell-1} \text{ and } \bar{x}_\ell \text{ solve}$$
$$\ddots$$
$$\text{where } \bar{x}_\ell \text{ solve}$$
$$\min_{\bar{x}_\ell} \gamma_1 \mathbb{E}_\omega[\bar{F}_\ell(\bar{x}_1, \ldots, \bar{x}_\ell, \omega)] + \gamma_2 \mathrm{Var}_\omega[\bar{F}_\ell(\bar{x}_1, \ldots, \bar{x}_\ell, \omega)] \quad (17)$$
$$\text{subject to } \mathrm{P}\{_\ell(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_\ell, \omega) \leq 0\} \geq \alpha,$$

$\bar{x}_1, \ldots, \bar{x}_{\ell-1}$ and $\bar{x}_\ell$ are decision variables expected to be decided before randomness is observed, $\alpha \in [0, 1]$ is reliability level, γ_1 and γ_2 are convex weight factors determining which one to minimize the most.

Problem (17) cannot be optimized directly due to the variance non-convexity property which is involved in the problem. Problem (17) can be changed into expectation criteria and the constraint becomes an a.s. constraint if the problem is risk-free and $\alpha = 1$. In such cases, the designed solution procedure for the TS-SMLP problem can be implemented also for SMLP risk-free with a.s. constraint type problem if F_i, G_i and x_i are set to be identically zero for all i and the assumptions presented in Theorem 2 hold. If problem (17) is non-risk free and/or $\alpha \in (0, 1)$, then problem (17a) can be mathematically rewritten as

$$\begin{aligned} &\min_{\bar{x}_i} \gamma_1 \frac{1}{n} \sum_{i=1}^{n} \bar{F}_i(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_\ell, \omega^i) + \\ &\quad \gamma_2 \frac{1}{n} \sum_{i=1}^{n} [\bar{F}_i(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_\ell, \omega^i) - \frac{1}{n} \sum_{i=1}^{n} \bar{F}_i(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_\ell, \omega^i)]^2 \\ &\text{subject to } \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}_{\bar{G}_i(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_\ell, \omega^i) \leq 0}[\bar{G}_i(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_\ell, \omega^i)] \geq \alpha, \ \forall \omega^i \end{aligned} \tag{18}$$

where $\mathbb{I}_{\bar{G}_i(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_i, \ldots, \bar{x}_\ell, \omega^i) \leq 0}$ is the indicator function and $\frac{1}{n} \sum_{i=1}^{n} [\bar{F}_i(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_i, \ldots, \bar{x}_\ell, \omega^i)$ $- \frac{1}{n} \sum_{i=1}^{n} \bar{F}_2(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_i, \ldots, \bar{x}_\ell, \omega^i)]^2$ approximates $\mathrm{Var}_\omega[\bar{F}_i(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_i, \ldots, \bar{x}_\ell, \omega)]$. In similar argument, rewrite all hierarchical level problems in problem (17) in the form of problem (18). Then, Algorithm 1 can be implemented on it to find a solution.

Any deterministic multilevel problem can be solved by the same solution procedure as long as it fulfills the deterministic versions of assumptions $\mathbf{A}_3$–$\mathbf{A}_6$.

4. Simulation results

Deterministic version problems from literature and some carefully constructed stochastic versions are given in subsection 4.1. The performance and total elapsed time of SSE solution procedure presents subsection 4.2.

4.1. SMLP test problems

Selected problems from literature, i.e., Examples 1–5, and constructed problems, i.e., Examples 6–8 are presented in this subsection, where the random variables which are involved in the problems are assumed to be either normally (N) or uniformly (U) distributed. Those test problems are used to show the effectiveness and competitiveness of the proposed solution procedure in the next subsection.

Example 1. The following trilevel problem is taken from [41].

$$\begin{aligned}
&\min_{x_1} \ -x_1 + 4x_2 \\
&\text{subject to } x_1 + x_2 \leq 1, \text{ where } x_2 \text{ and } x_3 \text{ solve} \\
&\qquad \min_{x_2} \ 2x_2 + x_3 \\
&\qquad \text{subject to } -2x_1 + x_2 \leq x_3, \text{ where } x_3 \text{ solve} \\
&\qquad\qquad \min_{x_3} \ -x_3^2 + x_2 \\
&\qquad\qquad \text{subject to } x_3 \leq x_1 \\
&\qquad\qquad\quad x_1 \in X_1 = [0, 0.5], x_2 \in X_2 = [0, 1] \\
&\qquad\qquad\quad x_3 \in X_3 = [0, 1]
\end{aligned}$$

Example 2. The following trilevel problem is taken from [60].

$$\begin{aligned}
&\min_{x_1} \ x_1^2 + 4x_2^2 + \sin^2(x_2 + x_3) - 6 \\
&\text{subject to } 3x_1 - 2x_2 - x_3 \leq 0, \ 2x_1 - x_2^2 + x_3^3 \leq 0, \\
&\qquad 2|x_1| - 3x_2 \leq 2, \text{ where } x_2 \text{ and } x_3 \text{ solve} \\
&\qquad \min_{x_2} \ x_1^2 + \frac{1}{5}\sin^2 x_2 \\
&\qquad \text{subject to } x_1 - x_2^2 \leq -x_3, \text{ where } x_3 \text{ solve} \\
&\qquad\qquad \min_{x_3} \ x_2^2 + x_3^2 \\
&\qquad\qquad \text{subject to } x_1 + x_2 \leq x_3 \\
&\qquad\qquad\quad x_1 \in X_1 = [-2, 2], x_2 \in X_2 = [-2, 2] \\
&\qquad\qquad\quad x_3 \in X_3 = [-2, 2]
\end{aligned}$$

Example 3. The following quad-level problem is taken from [30].

$$\begin{aligned}
&\min_{x_1} \ x_1^2 + 4x_2 - 2x_3 + x_4 \\
&\qquad \text{where } x_2, x_3 \text{ and } x_4 \text{ solves} \\
&\qquad \min_{x_2} \ 7x_1 - x_2^2 + 21x_3 - 2x_4 \\
&\qquad\qquad \text{where } x_3 \text{ and } x_4 \text{ solves} \\
&\qquad\qquad \min_{x_3} \ -x_1 + 7x_2 + x_3^2 - x_4^2
\end{aligned}$$

where x_4 solves

$$\min_{x_4} -x_1 + 3x_2 + 2x_1x_3 - 3x_4^2$$

subject to $\mathrm{x}_1 - 3x_2 + x_3^2 + x_4 \leq 32,$

$$-3\mathrm{x}_1 + 5x_2x_3 - x_3 - x_4 \leq 101,$$

$$3\mathrm{x}_1^2 + 5x_2 - x_3 + 2x_4 \leq 168$$

$$\mathrm{x}_1 \in X_1 = [0, 10], x_2 \in X_2 = [0, 5],$$

$$\mathrm{x}_3 \in X_3 = [0, 6], x_4 \in X_4 = [0, 3]$$

Example 4. The following quad-level problem is taken from [30].

$$\min_{x_1} -x_1 - 4x_2^2 - 2x_1x_3 - x_4$$

where x_2, x_3 and x_4 solve

$$\min_{x_2} -x_1x_3 - x_2 - x_3 - x_1x_4$$

where x_3 and x_4 solve

$$\min_{x_3} -x_1x_2^2 + 2x_2 - 2x_3^2 + x_2x_4$$

where x_4 solves

$$\min_{x_4} -x_1x_2 + x_2 - 3x_3 - x_3x_4$$

subject to $-x_1 - x_2 \leq -3,$

$$3x_1x_2^2 - 2x_2 + x_3 + x_4 \leq 10,$$

$$-2x_1 + x_2 - 2x_3 - x_4 \leq -1$$

$$x_1 \in X_1 = [0, 2], x_2 \in X_2 = [0, 8],$$

$$x_3 \in X_3 = [0, 8], x_4 \in X_4 = [0, 6]$$

Example 5. The following penta-level problem is taken from [34].

$$\min_{x_1} -7x_1^2 - x_2$$

where x_2, x_3,x_4, x_5 and x_6 solve

$$\min_{x_2,x_3} \cos(x_2)\mathrm{e}^{x_3} + x_1\mathrm{e}^{x_2x_3} + x_4$$

subject to $x_1 + x_2 - x_3 + x_6 - 5 \leq 0,$

where x_4, x_5 and x_6 solve

$$\min_{x_4} 2x_4^4 - x_1x_2 + x_3x_4 + x_6$$

subject to $-x_4 + x_2 + x_1 - 3 \leq 0$

where x_5 and x_6 solve

$$\min_{x_5} 5x_5 + x_1x_5^2 + x_3x_4x_5 + x_6$$

subject to $-2x_5 - x_4 - 3x_1 \leq 0$

where x_6 solve

$$\min_{x_6} x_6^2 - 8x_6 + x_5^2 + x_1^2 x_2^2$$
$$\text{subject to } x_6 + x_5 - 6 \leq 0, 2x_6 + x_4 - x_3 - 4 \leq 0$$
$$x_1 \in X_1 = [0,2], x_2 \in X_2 = [0,1], x_3 \in X_3 = [0,2],$$
$$x_4 \in X_4 = [0,1], x_5 \in X_5 = [0,2], x_6 \in X_6 = [0,1]$$

Example 6. The following example is a two stage stochastic trilevel programming problem in normal (N) probability distribution.

$$\min_{x_1,\bar{x}_1} -x_1 + 4x_2 + \mathbb{E}[2\omega\bar{x}_1^2 + (\bar{x}_2+\omega)^2 + \bar{x}_3]$$
$$\text{subject to } x_1 - x_2 \leq 25, x_1 + \bar{x}_1 - \bar{x}_3 - \omega \leq 20,$$
$$\text{where } x_2, \bar{x}_2, x_3 \text{ and } \bar{x}_3 \text{ solve}$$
$$\min_{x_2,\bar{x}_2} 2x_2 + x_3 + \mathbb{E}[\omega\bar{x}_1 + (\bar{x}_2+\omega)^2]$$
$$\text{subject to } 2x_1 - 2x_2 \leq x_3, x_1^2 + \bar{x}_1 - \bar{x}_3 + \omega\bar{x}_2^2 \leq 20,$$
$$\text{where } x_3 \text{ and } \bar{x}_3 \text{ solve}$$
$$\min_{x_3,\bar{x}_3} -x_3^2 + x_2 + \mathbb{E}[\omega\bar{x}_1 + \bar{x}_2^2 + \bar{x}_3]$$
$$\text{subject to } x_3 \leq x_1, -\bar{x}_2 - x_2\bar{x}_3 + \omega\bar{x}_2^2 \leq 0$$
$$x_1 \in X_1 = [0,2], x_2 \in X_2 = [0,1], x_3 \in X_3 = [0,1]$$
$$\bar{x_1} \in \bar{X}_1 = [0,6], \bar{x}_2 \in \bar{X}_2 = [0,5], \bar{x}_3 \in \bar{X}_3 = [0,4]$$
$$\forall\omega \in \Omega, \omega \sim N(2,3)$$

Example 7. The following example is a stochastic quad-level programming problem with almost sure constraint in normal (N) probability distribution.

$$\min_{\bar{x}_1} \mathbb{E}[\omega\bar{x}_1 + (\bar{x}_3+\omega)^2 + \bar{x}_4^2]$$
$$\text{subject to } \bar{x}_1 - 2\omega\bar{x}_3 + 2\bar{x}_4 \leq 50,$$
$$\text{where } \bar{x}_2, \bar{x}_3 \text{ and } \bar{x}_4 \text{ solve}$$
$$\min_{\bar{x}_2} \mathbb{E}[\bar{x}_1^2 + \omega\bar{x}_2^2 + \bar{x}_3\bar{x}_4]$$
$$\text{subject to } 3\omega\bar{x}_1 + \omega\bar{x}_2 - \omega\bar{x}_4^2 \leq 15,$$
$$\text{where } \bar{x}_3 \text{ and } \bar{x}_4 \text{ solve}$$
$$\min_{\bar{x}_3} \mathbb{E}[\bar{x}_2 + 2\omega\bar{x}_3^2 + \omega\bar{x}_4]$$
$$\text{subject to } \bar{x}_1^2 + 2\omega\bar{x}_2^2 + \omega\bar{x}_3 \leq 20$$
$$\text{where } \bar{x}_4 \text{ solves}$$
$$\min_{\bar{x}_4} \mathbb{E}[\bar{x}_1^2 - 2\bar{x}_2^2 + 2\omega\bar{x}_4]$$
$$\text{subject to } \bar{x}_2^3 + \bar{x}_2 - 2\bar{x}_3^2 + \omega\bar{x}_4^2 \leq 10$$
$$\bar{x}_1 \in \bar{X}_1 = [0,5], \bar{x}_2 \in \bar{X}_2 = [0,4]$$
$$\bar{x}_3 \in \bar{X}_3 = [0,6], \bar{x}_4 \in \bar{X}_4 = [0,8]$$
$$\forall\omega \in \Omega, \omega \sim N(2,3)$$

Example 8. The following example is a chance constraint penta-level programming in normal (N) probability distribution.

$$
\begin{aligned}
&\min_{x_1} \ \gamma_1\mathbb{E}[\omega\bar{x}_1^2 + 2\bar{x}_2^2 + \bar{x}_4 + 2\bar{x}_5] + \gamma_2 Var[\omega\bar{x}_1^2 + 2\bar{x}_2^2 + \bar{x}_4 + 2\bar{x}_5]\\
&\text{subject to } P\{2\omega\bar{x}_1^2 - 3\bar{x}_2^2 + \bar{x}_5 \leq 50\} \geq \alpha, \text{ where } \bar{x}_2, \bar{x}_3, \bar{x}_4 \text{ and } \bar{x}_5 \text{ solve}\\
&\quad \min_{\bar{x}_2} \ \gamma_1\mathbb{E}[\bar{x}_1^3 + \omega\bar{x}_2^2 + \bar{x}_3^2 - \bar{x}_4] + \gamma_2 Var[\bar{x}_1^3 + \omega\bar{x}_2^2 + \bar{x}_3^2 - \bar{x}_4]\\
&\quad \text{subject to } P\{3\bar{x}_1 + 5\bar{x}_2^2 + 6\omega\bar{x}_3 - \bar{x}_5 \leq 60\} \geq \alpha, , \text{ where } \bar{x}_3, \bar{x}_4 \text{ and } \bar{x}_5 \text{ solve}\\
&\qquad \min_{\bar{x}_3} \ \gamma_1\mathbb{E}[\bar{x}_1 + 2\omega\bar{x}_2^2 - \bar{x}_3 + \bar{x}_5^2] + \gamma_2 Var[\bar{x}_1 + 2\omega\bar{x}_2^2 - \bar{x}_3 + \bar{x}_5^2]\\
&\qquad \text{subject to } P\{3\omega\bar{x}_2^2 + 3\bar{x}_4^2 - \bar{x}_5 \leq 40\} \geq \alpha, \text{ where } \bar{x}_4 \text{ and } \bar{x}_5 \text{ solve}\\
&\qquad\quad \min_{\bar{x}_4} \ \gamma_1\mathbb{E}[\bar{x}_1^2 + x_3^2 + \omega\bar{x}_4^2 + \bar{x}_5] + \gamma_2 Var[\bar{x}_1^2 + \bar{x}_3^2 + \omega\bar{x}_4^2 + \bar{x}_5]\\
&\qquad\quad \text{subject to } P\{2\bar{x}_1 + 3\bar{x}_4^2 + 6\omega\bar{x}_5 \leq 30\} \geq \alpha, \text{ where } \bar{x}_5 \text{ solve}\\
&\qquad\qquad \min_{\bar{x}_5} \ \gamma_1\mathbb{E}[\bar{x}_1^2 - \bar{x}_3 + \bar{x}_4 + \omega\bar{x}_5^2] + \gamma_2 Var[\bar{x}_1^2 - \bar{x}_3 + \bar{x}_4 + \omega\bar{x}_5^2]\\
&\qquad\qquad \text{subject to } P\{2\bar{x}_1 + 5\bar{x}_3^2 + 6\omega\bar{x}_5 \leq 40\} \geq \alpha\\
&\qquad\qquad\quad \bar{x}_1 \in \bar{X}_1 = [2,6],\ \bar{x}_2 \in \bar{X}_2 = [-5,5],\ \bar{x}_3 \in \bar{X}_3 = [-5,5],\\
&\qquad\qquad\quad \bar{x}_4 \in \bar{X}_4 = [2,6],\ \bar{x}_5 \in \bar{X}_5 = [-5,5],\\
&\qquad\qquad\quad \gamma_1 = 0.6,\ \gamma_2 = 0.4,\ \alpha = 0.98,\ \forall\omega \in \Omega,\ \omega \sim N(2,1)
\end{aligned}
$$

In the next subsection, the simulation results of each test problem are discussed to show the effectiveness and competitiveness of the proposed SSE solution procedure.

4.2. Solution results and analysis

The simulation for the SSE solution procedure is performed using MATLAB 9.4.0(R2018a) software on an Intel core i3-380M laptop machine. The solution procedure is let to run 5 independent times and the best approximated solution is recorded for each simulation case.

First, we have to show the effectiveness and competitiveness of our proposed method for solving deterministic problems. Parameter values $\beta = 0.1$, $\beta = 0.5$, $\beta = 3$, $\beta = 1$ and $\beta = 1$ are set for each of the five problems in this category. Moreover, $\rho = 50$, $q = 5$ and $\pi = 10$ are set in the implementation of the solution procedure. The results from SSE method are comparedother methods' best results and presented in Table 2. In Examples 1, 2 and 4, the results obtained from the SSE method using equation (11) are superior in the objective values of all levels, whereas in Examples 3 and 5 the solutions obtained using the SSE solution procedure are better for some of the objectives (especially in the leader's objectives) and lag behind in some of the objectives of other levels. To closely analyze the total improvement of the result, percentage improvement (or variations) of our method for each of the decision makers in each of the problems is presented in Table 3. Here, the percentage improvement is computed using the formula

$$
\text{Percentage improvement} = \frac{\text{(New value – Original value)} \times 100\%}{\text{Original value}}. \tag{13}
$$

Table 2. Comparison of the results

No.	Method	Best solution	Best F_1	Best F_2	Best F_3	Best F_4	Best F_5
1	[60]	$(0.4994, 0.0016, 0.4988)$	-0.4929	0.5020	-0.2472		
	SSE	$(0.5, 0, 0.49951)$	-0.5	0.49951	-0.24951		
2	[60]	$(-0.0290, 0.0062, 0.0041)$	-5.9989	$8.4293e^{-004}$	$5.4521e^{-005}$		
	SSE	$(-0.0066, -3.5689e^{-06}, -1.275\ e^{-05})$	-6	$4.3969e^{-05}$	$1.7529e^{-10}$		
3	[30]	$(7.01, 4, 5.22, 2.53)$	57.23	137.63	11.82	58.97	
	SSE	$(0.02, 5, 0, 3)$	23	-30.8547	25.9792	-12.0208	
4	[30]	$(0, 5.51, 5.02, 4.38)$	-125.82	-10.53	-15.24	-31.53	
	SSE	$(0.10054, 6, 6, 5)$	-150.3071	-13.106	-33.6196	-42.6033	
5	[34]	$(2, 0.0398, 0.782, 0.999, 1.804, 1)$	-24.9604	5.2467	-5.0796	17.9474	-3.737
	SSE	$(2, 1, 0, 0, 0, 1)$	-29	2.5403	-1	1	-3

In Table 3, a positive percentage improvement value is assigned for decreased optimal value since the problem is minimization and a negative percentage improvement value is assigned for increased optimal value. All decision makers except the 3rd level decision maker's objective in Example 3 and, 3rd and 5rd levels objective values in Example 5 have shown improvements. In general, the solutions using SSE solution procedure (as can be seen from the average percentage performance values in the last column of Table 3) have shown improvement. Therefore, using the results in Table 2 and Table 3, the outcome of the new solution procedure is very much promising and competitive in terms of obtaining a good solution.

Table 3. Comparison results with SSE method [%]

No.	Comparison with SSE	Best F_1	Best F_2	Best F_3	Best F_4	Best F_5	Average improvement
1	[60]	$1,44$	0.49	0.93			0.95
2	[60]	0.018	94.8	99.9			64.9
3	[30]	59.8	122.4	-119.79	120.4		45.7
4	[30]	19.46	24.46	120.61	35.11		49.9
5	[34]	16.18	51.58	-80.31	94.42	-19.72	12.43

In addition, the quality of the solution will be improved by decreasing β as it is shown in Theorem 1. Now, to investigate the impact of increasing the value of β by some values in the performance of the algorithm, we demonstrate the results in Table 4 for deterministic case and in Table 7 for stochastic case. For the deterministic version, the performance of the method for two different values of β is presented in Table 4, where a parameter value $\rho = 50$, $q = 5$ and $\pi = 10$ are set in the code.

Using the results in Table 4, the improvement in the objective values of each level for two β values is considered so that one can see the cost and benefit of varying the β values in the algorithm. It can be seen from Table 4 that decision makers, except 3rd level's of Example 1 and, 3rd and $4th$ levels' of Example 3 are beneficial when the value of β is decreased. In addition to this, the time needed to solve the problem increases when the value of β decreases as presented in Table 4. So, we have to compromise between accuracy of the solution and the time it takes for the method to obtain the solution of the problem when choosing the value of β. The percentage improvement of each decision maker is presented in Table 5 to visualize the relation between improvement of optimal value of each decision maker and the variation of β values (as shown in the second column of the table).

Table 4. Two β parameter values comparison on each deterministic category problem

No.	β	Optimal solution	Best F_1	Best F_2	Best F_3	Best F_4	Best F_5	Time [s]
1	0.2	$(0.5, 0, 0.49954)$	-0.5	0.49983	-0.24983			16.1
	0.1	$(0.5, 0, 0.49951)$	-0.5	0.49951	-0.24951			39.9
2	0.75	$(-0.016, -1.218\text{e}^{-05}$, $7.757\ \text{e}^{-06})$	-5.9997	2.669e^{-004}	2.086e^{-10}			22.8
	0.5	$(-0.0066, -3.56\text{e}^{-06}$, -1.27 $\text{e}^{-05})$	-6	4.3969e^{-05}	1.7529e^{-10}			37
3	3	$(0.02, 5, 0, 3)$	23	-30.8547	25.9792	-12.0208		36.5
	1.5	$(0.000546, 5, 0, 3)$	23	-30.9962	25.9995	-12.0005		134
4	2	$(0.0695, 6, 6, 5)$	-149.9036	-12.7646	-32.5022	-42.417		46.7
	1	$(0.10054, 6, 6, 5)$	-150.3071	-13.106	-33.6196	-42.6033		84.4
5	1	$(2, 1, 0, 0, 0, 1)$	-29	2.5403	-1	1	-3	107
	0.5	$(2, 1, 0, 0, 0, 1)$	-29	2.5403	-1	1	-3	434

As can be seen in Table 5, as β is decreased by some value as shown in the second column of the table, the objective values of all decision makers except 3rd level's of Example 1 and, 3rd and $4th$ levels' of Example 3 have shown improvement. However, the solutions using smaller β values (as can be seen from the $8th$ column in Table 5) are good in terms of the average percentage improvement values except for Example 1. The average percentage improvement value for Example 1 shown positive due to the larger improvement of the 3rd level decision maker's optimal value compared to its 2nd level decision maker's optimal value which is poor. However, the time needed to run the program for each problem is increased (cf. the last column of Table 5).

Table 5. Two different β parameter value percentage comparison results for each problem [%]

No.	Change of β	Best F_1	Best F_2	Best F_3	Best F_4	Best F_5	Average improvement	Time [s]
1	from 0.2 to 0.1	0	0.064	-0.128			-0.02	23.8
2	from 0.75 to 0.5	0.005	83.52	16.15			33.225	13.7
3	from 3 to 1.5	0	0.458	-0.07	-0.168		0.055	97.5
4	from 2 to 1	0.269	2.67	3.43	0.439		1.702	37.5s
5	from 1 to 0.5	0	0	0	0	0	0	327s

The performance of the solution procedure is highly dependent on the value of β which is the parameter to be chosen. The simulation results can be improved further by choosing smaller values for β but the time needed for solving the problem will be increased (cf. last column of Tables 4 and 5). For a stochastic version case, optimal solution and optimal values are presented in Table 6, where parameter values $\rho = 50$, $\pi = 10$, $q = 5$ and $n = 25$ are set in the code, are the results by the SSE method.

Table 6. Solutions by the SSE method

No.	β	Optimal solution	Best $\mathbb{F}_1$	Best $\mathbb{F}_2$	Best $\mathbb{F}_3$	Best $\mathbb{F}_4$	Best $\mathbb{F}_5$
6	0.75	$(0.19662, 0.19585, 0.040776, 0, 0, 0)$	6.1185	6.3667	0.25955		
7	1.5	$(0.020759, 0, 0, 0)$	10.3848	0.00043093	0	0.00043093	
8	2	$(2.0048, 0.0026364, 2.6926, 2, -0.29469)$	9.8227	7.9851	0.36056	15.4954	2.0906

Like in the deterministic case, the impact of increasing β by some values in the performance of the solution procedure is demonstrated in Table 7, where $\rho = 50$, $\pi = 10$, $q = 5$ and $n = 25$ are set in the code.

Table 7. Two different β parameter value comparison results on each SMLP problem

No.	β	Optimal solution	Best $\mathbb{F}_1$	Best $\mathbb{F}_2$	Best $\mathbb{F}_3$	Best $\mathbb{F}_4$	Best $\mathbb{F}_5$	Time [s]
6	1	$(0.24226, 0.2463, 0.07463, 0, 0, 0)$	9.7359	9.9587	0.33044			491
	0.75	$(0.19662, 0.19585,$ 0.040776, 0,0,0)	6.1185	6.3667	0.25955			1621
7	2	$(0.036024, 0, 0, 0)$	12.6529	0.0012978	0	0.0012978		266
	1.5	$(0.020759, 0, 0, 0)$	10.3848	0.00043093	0	0.00043093		424
8	2.5	$(2.104, -0.0253, 2.723, 2.03, -1.212)$	12.0084	8.824	0.51236	17.2929	4.6429	1217
	2	$(2.0048, 0.00263,$ 2.6926,2, -0.29469)	9.8227	7.9851	0.36056	15.4954	2.0906	2258

It can be seen from Table 7 that all optimal values shown are with high quality when the value of β is decreased (cf. the 2nd column of the table). The improvement in the payoff values for each decision maker in each problem and the average improvement result when the value of β is decreased are analyzed in Table 8. As β is decreased by the given value (cf. the second column of the table), the optimal values for each decision maker in all the problems become better. The solutions using smaller β values (cf. the average percentage improvement column in Table 8) are improving. In general, the results can be improved further by choosing smaller values for β with the parameter values of n kept fixed. However, in this case, the time elapsed for solving the problem will be increased (cf. the last column of Tables 7 and 8) like in the case of the deterministic version.

Table 8. Two different β parameter value percentage comparison results on each SMLP problem

No.	Change of β	Best $\mathbb{F}_1$	Best $\mathbb{F}_2$	Best $\mathbb{F}_3$	Best $\mathbb{F}_4$	Best $\mathbb{F}_5$	Average improvement	Time increment [s]
6	from 1 to 0.75	37.1	36	21.4			31.5	1130
7	from 2 to 1.5	17.9	66.8	0	66.8		37.8	158
8	from 2.5 to 2	18.2	9.5	29.6	10.4	54.9	24.5	1041

Table 9. Two different n parameter value comparison results on each SMLP problem

No.	n	Optimal solution	Best $\mathbb{F}_1$	Best $\mathbb{F}_2$	Best $\mathbb{F}_3$	Best $\mathbb{F}_4$	Best $\mathbb{F}_5$	Time [s]	SSE parameters
6	25	$(0.24226, 0.2463,$ 0.07463,0, 0, 0)	9.7359	9.9587	0.33044			491	$\beta = 1$ $\rho = 50,$
	50	$(0.17629, 0.17457,$ 0.11484, 0.044437, 0.28521, 0)	11.9612	12.3577	0.28851			904	$\pi = 10,$ $q = 5$
7	25	$(0.036024, 0, 0, 0)$	12.6529	0.0012978	0	0.0012978		266	$\beta = 2,$ $\rho = 50,$
	50	$(0.028274, 0, 0, 0)$	10.8594	0.00079944	0	0.00079944		511	$\pi = 10,$ $q = 5$
8	25	$(2.104, -0.0253,$ 2.723, 2.03, $-1.212)$	12.0084	8.824	0.51236	17.2929	4.6429	1217	$\beta = 2.5,$ $\rho = 50,$
	50	$(2.0054, -1.6607,$ $4.9858, 2.0567,$ $-4.9832)$	9.7303	24.806	31.6684	26.4292	272.9188	2086	$\pi = 10,$ $q = 5$

In addition to the parameter β, the choice of the value of n for stochastic version problems also affects the solution quality and the time needed for solving the problem. As n increases, the number of functional evaluations and the number of constraints, which are formed by the number of realization n, increases.

Table 9 presents the optimal values and the time needed for solving the SMLP problem for two different values of n. It can be seen from Table 9 that the objective function value of the 3rd decision maker in Example 6, all decision-maker's objective values in Example 7 and the 1st level decision maker's in Example 8 have shown improvement when the value of n is increased (as indicated in the 2nd column of the table). The improvement in the payoff values for each decision maker in each problem and the average improvement result, when the value of n is increased, are analyzed in Table 10. As n is increased by the given value (cf. the second column of the table), the optimal values for the 3rd level decision maker in Example 6, all decision-makers in Example 7 and the 1st level decision maker in Example 8 show improvement. But, only the performance in Example 7 has produced a good result in terms of average percentage improvement. Notice that as $n \to \infty$, sample average approximation converges under suitable conditions (like convexity condition) but some of the decision makers in Examples 6 and 8 have non-convexity property. In addition, the convergence to an optimal solution for one of the decision makers in the hierarchy can affect the optimal values of the other decision makers.

Table 10. Two different n parameter value percentage comparison results on each SMLP problem [%]

No.	Change of n	Best $\mathbb{F}_1$	Best $\mathbb{F}_2$	Best $\mathbb{F}_3$	Best $\mathbb{F}_4$	Best $\mathbb{F}_5$	Average improvement	Time increment [s]
6	from 25 to 50	-22.8	-24	12.68			-11.3	413
7	from 25 to 50	14.1	38.4	0	38.4		22.725	245
8	from 25 to 50	18.9	-124.1	-6085	-52.8	-5778	-1187	869

Therefore, the proposed SSE method is very promising in finding approximate solutions for hierarchical problems and is also competitive in solving various types of SMLP problems especially when in the implementation the parameter value for β is taken to be sufficiently small. In addition, if a larger parameter value for π, q, and/or n is set, the time required for the algorithm to solve the problem increases but the approximated solution for each decision makers will show improvement.

5. Conclusion

In this study, stochastic multilevel programming (SMLP) problems have been considered. The SSE method, which is a non-derivative meta-heuristic type algorithm, is proposed for solving the problem. The algorithm is constructed based on the realization of the random variables to convert the problems into deterministic form. Then the procedure employs systematic partitioning of each level's decision space to search for an optimal reaction from each decision space starting from the last level problem up through the first level problem. Meanwhile, the PSO algorithm is applied to find the best response for the decision of each hierarchical level. The process of updating responses for each leader's action continues iteratively until the maximum number of iterations is attained.

The existence of a solution and convergence of the algorithm is established. Since stochastic optimization methods generalize deterministic methods, the proposed algorithm is compared with other methods from the literature on a deterministic version of the problem. The results of the numerical simulations are very much promising. The solution procedure can be used to solve complex stochastic multilevel programming and deterministic multilevel programming problems. It can also be considered as a new additional topic for stochastic multilevel programming problems.

The proposed method can handle multilevel problems of any type ranging from continuous optimization problems with any type of constraint to problems with mixed-integer decision variables. Since it uses meta-heuristic methods, smoothness of the involved functions is not required. In addition, since a penalty method is implemented to discriminate a non-feasible solution at each stage, the method guarantees the multilevel feasibility of the solution obtained in each iteration. Moreover, the use of PSO in solving the most inner-level problem allows to obtain an approximate global optimal solution even if the problem is not convex.

However, since the method uses an inductive combination of selected points from each decision space, the time the algorithm requires to solve a given problem increases with the increase in dimension, hierarchical levels, and number of partitions of the decision spaces (i.e., when β is small). The possibility of reducing the running time of the algorithm in these cases might be one of the topics for future investigation. As in any other meta-heuristic method, one may not get the same solution when the algorithm is repeated. However, the quality of the solution can be controlled by the algorithmic parameter β. In addition, the realization of the random variables is an area where it requires improvement in future studies.

Acknowledgement

The authors are grateful to the anonymous reviewers of the journal for their constructive and insightful comments, which have improved the earlier version of the manuscript.

References

[1] Acevedo, J., and Pistikopoulos, E. N. A multiparametric programming approach for linear process engineering problems under uncertainty. *Industrial Engineering Chemistry Research 36*, 3 (1997), 717–728.

[2] Ahmadi-Moshkenani, P., Johansen, T. A., and Olaru, S. Combinatorial approach towards multi-parametric quadratic programming based on characterizing adjacent critical regions. *IEEE Transactions on Automatic Control 63*, 10 (2018), 3221–3231.

[3] Alemdar, N. M., and Sirakaya, S. Online computation of Stackelberg equilibria with synchronous parallel genetic algorithms. *Journal of Economic Dynamics and Control 27*, 8 (2003), 1503–1515.

[4] Alizadeh, S. M., Marcotte, P., and Savard, G. Two stage stochastic bilevel programming over a transportation network. *Transportation Research Part B: Methodological 58* (2013), 92–105.

[5] Audestad, J.-A., Gaivoronski, A. A., and Werner, S. Extending the stochastic programming framework for the modeling of several decision makers: pricing and competition in the telecommunication sector. *Annals of Operations Research 142* (2006), 19–39.

[6] Avraamidou, S., and Pistikopoulos, E. N. Multi-parametric global optimization approach for trilevel mixed-integer linear optimization problems. *Journal of Global Optimization 74* (2019), 443–465.

[7] Avraamidou, S., and Pistikopoulos, E. N. A multi-parametric optimization approach for bilevel mixed-integer linear and quadratic programming problems. *Computers & Chemical Engineering 125* (2019), 98–113.

[8] Awraris, A., Wordofa, B. G., and Kassa, S. M. A simplex-based branch-and-cut method for solving integer trilevel programming problems. *Journal of Mathematics and Computer Science 25* (2022), 232–250.

[9] Axehill, D., Besselmann, T., Raimondo, D. M., and Morari, M. A parametric branch and bound approach to suboptimal explicit hybrid MPC. *Automatica 50*, 1 (2014), 240–246.

[10] Bard, J. A grid search algorithm for the linear bilevel programming problem. In *Proceedings of the 14th Annual Meeting of the American Institute for Decision Science* (1982), pp. 256–258.

[11] Ben-Ayed, O. Bilevel linear programming. *Computer and Operations Research 20*, 5 (1993), 485–501.

[12] Beykal, B., Avraamidou, S., and Pistikopoulos, E. N. Bi-level mixed-integer data-driven optimization of integrated planning and scheduling problems. In *31st European Symposium on Computer Aided Process Engineering*, M. Türkay and R. Gani, Eds., vol. 50 of *Computer Aided Chemical Engineering*. Elsevier, 2021, pp. 1707–1713.

[13] Beykal, B., Avraamidou, S., and Pistikopoulos, E. N. Data-driven optimization of mixed-integer bi-level multi-follower integrated planning and scheduling problems under demand uncertainty. *Computers & Chemical Engineering 156* (2022), 107551.

[14] Beykal, B., Avraamidou, S., Pistikopoulos, I. P. E., Onel, M., and Pistikopoulos, E. N. DOMINO: Data-driven Optimization of bi-level Mixed-Integer NOnlinear Problems. *Journal of Global Optimization 78* (2020), 1–36.

[15] Bialas, W. F., and Karwan, M. K. *Mathematical Methods for Multilevel Planning*. Research Report No. 79-2, Department of Industrial Engineering, State university of New York at Buffalo, NY, 1979.

[16] Borrelli, F., Bemporad, A., and Morari, M. Geometric algorithm for multiparametric linear programming. *Journal of Optimization Theory and Applications 118*, 3 (2003), 515–540.

[17] Bracken, J., and McGill, J. T. Mathematical programs with optimization problems in the constraints. *Operations Research 21*, 1 (1973), 37–44.

[18] Bracken, J., and McGill, J. T. Technical note – A method for solving mathematical problems with nonlinear programs in the constraints. *Operations Research 22*, 5 (1974), 1097–110l.

[19] Christiansen, S., Patriksson, M., and Wynter, L. Stochastic bilevel programming in structural optimization. *Structural and Multidisciplinary Optimization 21*, 5 (2001), 361–371.

[20] Dempe, S. *Foundations of Bilevel Programming*. Kluver Academic Publisers, New York, NY, 2002.

[21] Dempe, S. Comment to "interactive fuzzy goal programming approach for bilevel programming problem" by S. R. Arora and R. Gupta. *European Journal of Operational Research 212*, 2 (2011), 429–431.

[22] Dempe, S. *Bilevel Optimization: Theory, Algorithms and Applications*. Preprint 2018-11, Fakultät für Mathematik und Informatik, TU Bergakademie Freiberg, 2018.

[23] Dempe, S., Ivanov, S., and Naumov, A. Reduction of the bilevel stochastic optimization problem with quantile objective function to a mixed-integer problem. *Applied Stochastic Models in Business Industry 33*, 5 (2017), 544–554.

[24] Faísca, N. P., Saraiva, P. M., Rustem, B., and Pistikopoulos, E. N. A multi-parametric programming approach for multilevel hierarchical and decentralised optimisation problems. *Computational Management Science 6*, 4 (2009), 377–397.

[25] Goshu, N. N., and Kassa, S. M. A systematic sampling evolutionary (SSE) method for stochastic bilevel programming problems. *Computers and Operations Research 120* (2020), 104942.

[26] Goshu, N. N., and Kassa, S. M. Stochastic bilevel programming with multiple followers: A solution approach using the systematic sampling evolutionary method. *Engineering Optimization 54*, 6 (2022), 1059–1072.

[27] Goshu, N. N., and Kassa, S. M. Stochastic sequential supply chain management system: with a solution approach using the systematic sampling evolutionary method. *International Journal of Business Performance and Supply Chain Modelling 13*, 3 (2022), 264–288.

[28] Han, J., Zhang, G., Hu, Y., and Lu, J. Solving trilevel programming problems using a particle swarm optimization algorithm. In *Proceedings of the 10th IEEE Conference on Industrial Electronics and Applications*, (Auckland, New Zealand), 2015, IEEE, pp. 569–574.

[29] Han, J., Zhang, G., Hu, Y., and Lu, J. A solution to bi/trilevel programming problems using particle swarm optimization. *Information Sciences 370–371* (2016), 519–537.

[30] Hosseini, E. Presentation and solving non-linear quad-level programming problem utilizing a heuristic approach based on Taylor theorem. *Journal of Optimization in Industrial Engineering 11*, 1 (2018), 91–101.

[31] Ivanov, S. A bilevel stochastic programming problem with random parameters in the follower's objective function. *Journal of Applied and Industrial Mathematics 12*, 4 (2018), 658–667 .

[32] Kassa, A. M., and Kassa, S. M. A multi-parametric programming algorithm for special classes of non-convex multilevel optimization problems. *An International Journal of Optimization and Control: Theories & Applications 3*, 2 (2013), 133–144.

[33] Kassa, A. M., and Kassa, S. M. Approximate solution algorithm for multi-parametric non-convex programming problems with polyhedral constraints. *An International Journal of Optimization and Control: Theories & Applications 4*, 2 (2014), 89–98.

[34] Kassa, A. M., and Kassa, S. M. A branch-and-bound multi-parametric programming approach for non-convex multilevel optimization with polyhedral constraints. *Journal of Global Optimization 64*, 4 (2016), 745–764.

[35] Kassa, A. M., and Kassa, S. M. Deterministic solution approach for some classes of nonlinear multilevel programs with multiple followers. *Journal of Global Optimization 68*, 4 (2017), 729–747.

[36] Kassa, S. M. Three-level global resource allocation model for HIV controle: A hierarchical decision system approach. *Mathematical Biosciences and Engineering 15*, 1 (2018), 255–273.

[37] Kassa, S. M., and Tsegay, T. H. An iterative method for trilevel quadratic fractional programming problems using fuzzy goal programming approach. *Journal of Industrial Engineering International 14*, 2 (2018), 255–264.

[38] Kim, S., Pasupathy, R., and Henderson, S. G. A Guide to Sample Average Approximation. In *Handbook of Simulation Optimization*, C. M. Fu, Ed., vol. 216 of International Series in Operations Research & Management Science, Springer, New York, NY, 2015, pp. 207–243.

[39] Kumar, M., and Pal, B. B. Fuzzy goal programming approach to chance constrained multilevel programming problems. *International Journal of Advanced Computer Research 3*, 8 (2013), 193–200.

[40] Lv, Y., and Jiang, J. A relaxation solving approach for the linear trilevel programming problem. *Computational and Applied Mathematics 40*, 6 (2021), 226.

[41] Molla, A. *A Multiparametric Programming Approach for Multilevel Optimization*. MSc thesis, Department of Mathematics, Addis Ababa University, 2011.

[42] Pappas, I., Diangelakis, N. A., and Pistikopoulos, E. N. The exact solution of multiparametric quadratically constrained quadratic programming problems. *Journal of Global Optimization 79*, 1 (2021), 59–85.

[43] Patriksson, M. On the applicability and solution of bilevel optimization models in transportation science: A study on the existence, stability and computation of optimal solutions to stochastic mathematical programs with equilibrium constraints. *Transportation*

Research Part B: Methodological 42, 10 (2008), 843–860.

[44] Patriksson, M., and Wynter, L. Stochastic mathematical programs with equilibrium constraints. *Operation Research Letters 25*, 4 (1999), 159–167.

[45] Pistikopoulos, E. N., Georgidis, M. C., and Dua, V. *Multi-Parametric Programming: Theory, Algorithms and Applications*. Vol. 1, Wiley-VCH, Weinheim, 2007.

[46] Pramanik, S., and Banerjee, D. Chance constrained quadratic bi-level programming problem. *International Journal of Modern Engineering Research 2*, 4 (2012), 2417–2424.

[47] Pramanik, S., Banerjee, D., and Giri, B. Chance constrained linear plus linear fractional bilevel programming problem. *International Journal of Computer Applications 56*, 16 (2012), 34–39.

[48] Pramanik, S., Banerjee, D., and Giri, B. Chance constrained multi-level linear programming problem. *International Journal of Computer Applications 120*, 18 (2015), 1–6.

[49] Robbins, H., and Monro, S. A stochastic approximation method. *The Annals of Mathematics Statistics 22*, 3 (1951), 400–407.

[50] Sakawa, M., Nishizaki, I., and Uemura, Y. Interactive fuzzy programming for multilevel linear programming problems. *Computers and Mathematics with Applications 36*, 2 (1998), 71–86.

[51] Shi, C., Zhang, G., and Lu, J. The Kth-best approach for linear bilevel multi-follower programming. *Journal of Global Optimization 33*, 4 (2005), 563–578.

[52] Smola, A. J., Vishwanathan, S. V. N., and Le, Q. V. Bundle methods for machine learning. In *Advances in Neural Information Processing Systems 20 (NIPS 2007)*, J. Platt, D. Koller, Y. Singer and S. Roweis, Eds., Curran Associates, Inc., 2008, pp. 1377–1384.

[53] von Stackelberg, H. *The Theory of the Market Economy*. Oxford University Press, New York, NY, 1952.

[54] Tilahun, S. L., Kassa, S. M., and Ong, H. C. A new algorithm for multilevel optimization problems using evolutionary strategy, inspired by natural adaptation. In *PRICAI 2012: Trends in Artificial Intelligence* , P. Anthony, M. Ishizuka, and D. Lukose, Eds., vol. 7458, (Berlin, Heidelberg, 2012), Springer, pp. 577–588.

[55] Vallée, T., and Başar, T. Off-line computation of Stackelberg solutions with the genetic algorithm. *Computational Economics 13*, 3 (1999), 201–209.

[56] Vicente, L. N., and Calamai, P. H. Bilevel and multilevel programming: A bibliography review. *Journal of Global Optimization 5*, 3 (1994), 291–306.

[57] Wang, Y., Jiao, Y.-C., and Li, H. An evolutionary algorithm for solving nonlinear bilevel programming based on a new constraint-handling scheme. *IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews 35*, 2 (2005), 221–232.

[58] Werner, S. *Bilevel Stochastic Programming Problems: Analysis and Application to Telecommunications*. PhD thesis, Norwegian University of Science and Technology, Trondheim 2004.

[59] White, D. J. Multilevel programming, rational reaction sets, and efficient solutions. *Journal of Optimization Theory and Applications 87*, 3 (1995), 727–746.

[60] Woldemariam, A. T., and Kassa, S. M. Systematic evolutionary algorithm for general multilevel Stackelberg problems with bounded decision variables (SEAMSP). *Annals of Operations Research 229*, 1 (2015), 771–790.

[61] Yeh, K., Whittaker, C., Realff, M. J., and Lee, J. H. Two stage stochastic bilevel programming model of a pre-established timberlands supply chain with biorefinery investment interests. *Computers and Chemical Engineering, Elsevier 73*, 2 (2015), 141–153.

[62] Zewde, A. B., and Kassa, S. M. A novel approach for solving multi-parametric problems with nonlinear constraints. *Journal of Global Optimization 85*, 2 (2023), 283–313.

[63] Zhan, Y., and Zheng, Q. P. A multistage decision-dependent stochastic bilevel programming approach for power generation investment expansion planning. *IISE Transactions 50*, 8 (2018), 720–734.

[64] Zhang, G., Lu, J., Montero, J., and Zeng, Y. Model, solution concept, and Kth-best algorithm for linear trilevel programming. *Information Sciences 180*, 4 (2010), 481–492.