ARTICLE IN PRESS

European Journal of Operational Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Decision Support

A simulation optimization approach for weight valuation in analytic hierarchy process

Hui Xiao [c,e], Sha Zeng [a], Yi Peng [b,*], Gang Kou [d,e]

[a] School of Business Administration, Southwestern University of Finance and Economics, Chengdu 611130, China
[b] School of Management and Economics, University of Electronic Science and Technology of China, Chengdu 611731, China
[c] School of Management Science and Engineering, Southwestern University of Finance and Economics, Chengdu 611130, China
[d] Xiangjiang Laboratory, Changsha 410205, China
[e] Big Data Laboratory on Financial Security and Behavior, SWUFE (Laboratory of Philosophy and Social Sciences, Ministry of Education), Chengdu 611130, China

ARTICLE INFO

Keywords:
Analytic hierarchy process (AHP)
Knowledge gradient (KG)
Multiple criteria analysis
Ranking and selection (R&S)

ABSTRACT

The analytic hierarchy process (AHP) is a structured technique used to analyze complex decision-making situations such as resource allocation, benchmarking, and quality management. In the weight valuation step of using AHP to select the best design, pairwise comparison matrices are used to calculate the local priorities for designs that have contentious and unresolved criticisms. In this study, we propose a Bayesian approach using a Dirichlet-multinomial model to estimate local priorities during weight valuation. Experts are only asked to select the best design with respect to predetermined criterion. Subsequently, local priorities are estimated without pairwise comparison matrices. To improve the efficiency of the AHP, we propose two expert allocation policies (AHP-KG and AHP-AKG) based on the ranking and selection procedures. Our numerical results show that the proposed AHP-KG and AHP-AKG policies outperform pure exploration and proportional allocation policies.

1. Introduction

The analytic hierarchy process (AHP) is one of the most frequently used decision support tool (Saaty, 1980; Brunelli, 2014; Ho & Ma, 2018). With its hierarchical structure, the AHP considers several factors simultaneously and arrives at a conclusion (Saaty & Vargas, 2012). This method has been used in many decision-making areas, such as probability assessment, selection, evaluation, benefit-cost analysis, planning and development, priority and ranking, medicine, and quality function deployment (Ho, 2008; Subramanian & Ramanathan, 2012; Kokangül et al., 2017; Ping et al., 2018; Lyu et al., 2020).

In its general form, the AHP comprises four major steps: problem modeling, weight valuation, weight aggregation, and sensitivity analysis (Ishizaka & Labib, 2011). The three-level hierarchy, consisting of decision goal, criteria, and designs, is the basic and simplest form of any hierarchical structure in AHP. Even hierarchical structures with more levels can be decomposed into groups of three-level hierarchies (Huang et al., 2015). Therefore, in the following sections, we focus on the three-level hierarchy.

Fig. 1 shows an example of how to use the AHP to select the best design for rough terrain cargo handlers (Bard & Sousk, 1990). In the weight valuation step, experts from different fields are invited to construct pairwise comparison matrices (PCMs) of the designs or criteria. Criterion weights, representing the weights of each criterion for the target, and local priorities, the weights of each design with respect to each criterion, are determined using PCMs. Next, global priorities, that is, the weights of designs with respect to the target, can be derived by aggregating the criterion weights and local priorities, and the best design can be selected based on global priorities.

In this study, we consider the problem of using the AHP to select the best design from a finite number of designs. Traditionally, PCM has been an important tool in the weight valuation step of the AHP. PCMs are derived from experts' subjective judgments, and the number of pairwise comparisons requested can be very high: $(K^2 - K)/$ for K designs. There are some criticisms associated with PCMs in terms of weight valuation. First, experts should evaluate the designs using the same measurement scales in order to construct the PCMs. The choice of the best scale is a topic of heated debate and it is difficult to verify the effectiveness of scales through subjective issues (Ishizaka & Labib, 2011). Second, the curse of dimensionality in PCMs, which results in inconsistency or incompleteness, limits their application in large-scale decision-making problems (Jalao et al., 2014). Furthermore, the computational efficiency

* Corresponding author.
E-mail address: pengyi@uestc.edu.cn (Y. Peng).

https://doi.org/10.1016/j.ejor.2024.10.018
Received 16 November 2022; Accepted 11 October 2024
Available online 12 October 2024

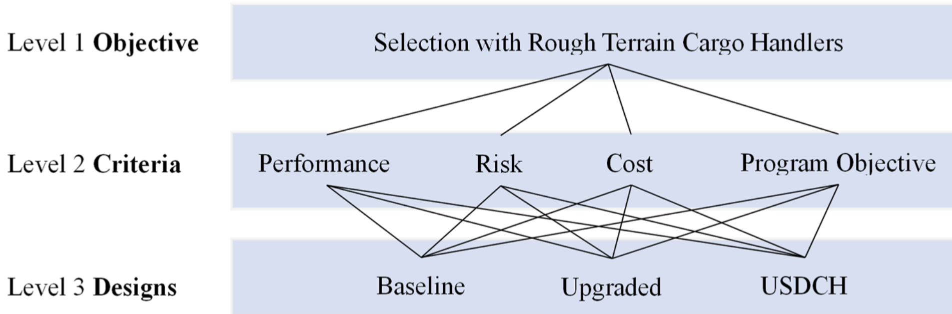

Fig. 1. Hierarchical Structure of Rough Terrain Cargo Handlers Selection Problem.

is a significant challenge when deriving priorities from large-scale sparse PCMs (Wang et al., 2021). Finally, owing to the limitations of human knowledge and experience, the evaluation results of experts often differ, requiring multiple experts to evaluate the design performance. The evaluations may be inconsistent or incomplete because of the limitations of the experts' capabilities or the complexity of the decision problems, which will lead to inconsistencies in the PCMs (Lin & Kou, 2015; Kou et al., 2016; Aguarón et al., 2021). While inviting more experts can improve the accuracy of estimating local priorities, it also increases the procedure cost. Therefore, deriving an expert allocation rule is essential for improving the efficiency of evaluation procedures.

Bayesian prioritization procedures are used to derive priorities from the PCMs in the weight valuation step. For example, several studies assume that local priorities satisfy log-normal distributions and use Bayesian analysis to estimate priorities in group decision making (Altuzarra et al., 2007; Gargallo et al., 2007). Lin and Kou (2015) propose a Bayesian method for revising individual PCMs before aggregating experts' evaluations. The expert allocation problem in the AHP can be formulated as a Ranking and Selection (R&S) problem. R&S problems involve deriving an efficient budget allocation rule, such that the selection quality can be optimized for a given fixed budget (Xiao et al., 2019). Examples of R&S problems include the selection of Pareto set (Lee et al., 2010; Li et al., 2018), robust R&S (Xiao & Gao, 2018; Fan et al., 2020), parallel computing (Wu & Frazier, 2016; Zhong & Hong, 2021), and search algorithms (Zhang et al., 2016; Kou et al., 2021). Huang et al. (2015) allocates experts within the AHP framework using the optimal computing budget allocation method, which is a popular R&S approach.

All the aforementioned studies are based on PCMs constructed by experts but do not address the problems associated with PCMs, such as measurement scales, inconsistency issues, and priority derivation for incomplete or large-scale PCMs.

To solve the problem of using AHP to select the best design, we estimate the local priorities from a Bayesian perspective using the Dirichlet-multinomial model, bypassing PCM issues. To improve the efficiency of the AHP, we formulate the expert allocation problem as an R&S problem and generalize the knowledge gradient (KG) policy (Frazier et al., 2008) and asynchronous knowledge gradient (AKG) policy (Kaminski & Szufel, 2014).

Experts are invited to select the best design with respect to one criterion, and the evaluation results are assumed to satisfy multinomial distributions with parameters equal to the local priorities. From a Bayesian perspective, the local priorities are interpreted as random variables with Dirichlet priors that are conjugate to the multinomial distributions. The parameters of the Dirichlet distributions are updated based on the evaluation results. Finally, we estimate the local priorities using the mean values of the Dirichlet distributions to determine the design with the highest global priority.

Based on the Dirichlet-multinomial model, the KG and AKG policies are generalized to solve the expert allocation problem in the weight valuation step of the AHP. The KG policy sequentially samples alternatives and optimizes the expected value of the information gained from a single additional sample (Groves & Branke, 2019). The AKG policy extends the KG policy by enabling parallel and asynchronous dispatch of evaluation tasks among workers, thereby improving the efficiency of the AHP.

The contributions of this study are twofold. First, it is the first study to consider the weight valuation step in AHP using the Dirichlet-multinomial model from a Bayesian perspective. Experts only need to select the best design with respect to one criterion, and local priorities can be estimated without using PCMs. Second, we formulate KG and AKG policies based on the Dirichlet-multinomial model to solve the expert allocation problem in AHP and derive two efficient budget allocation procedures for selecting the best design within the AHP.

We describe this problem in detail in the following sections. Section 2 formulates the problem, Sections 3 and 4 introduces the AHP-KG and AHP-AKG policies applied to the problem. Section 5 presents the results of numerical experiments, demonstrating a real case study and different scenarios. Finally, Section 6 concludes the paper.

2. Problem formulation

We consider the problem of using AHP to select the best design from a finite number of K designs. The performance of design k is represented by global priority g_k, which is unknown. The global priority g_k is the arithmetic mean of the local priorities $p_{k,l}$ for each criterion l, which are estimated from experts' evaluations (Ishizaka & Labib, 2011). Since experts' evaluations may be influenced by individual factors, inviting more experts can improve the accuracy of the estimation of local priorities. However, this increases the procedure cost. We assume a budget N that determines the number of times experts evaluate it, and we derive an expert allocation rule to improve the efficiency of the evaluation procedure.

To formulate the problem, we define the following notations.

L	the number of criteria in the second level of the hierarchy.
K	the number of designs.
N	the budget of the evaluation.
w_l	the weight of criterion l, and $l = 1, 2, \cdots, L$. $\boldsymbol{W} \triangleq (w_1, w_2, \cdots, w_L)^T$ denotes all criterion weights with respect to the objective, and $\boldsymbol{W} \in T_L(1)$. Herein, $T_L(1) = \left\{(w_1, w_2, \cdots, w_L)^T \middle| \sum_{l=1}^{L} w_l = 1, w_l > 0,\ \forall 1 \le l \le L\right\}$.
$p_{k,l}$	the local priority of design k with respect to criterion l, which is unknown. $\boldsymbol{p}_l \triangleq \left(p_{1,l}, p_{2,l}, \cdots, p_{K,l}\right)^T$ denotes all local priorities with respect to criterion l; herein, $\boldsymbol{p}_l \in T_K(1)$.
$\widehat{p}_{k,l}^n$	the estimation of the local priority $p_{k,l}$ at time $0 \le n \le N$, and $\widehat{\boldsymbol{P}}_l^n \triangleq \left(\widehat{p}_{1,l}^n, \widehat{p}_{2,l}^n, \cdots, \widehat{p}_{K,l}^n\right)^T$.
g_k	the global priority of design k. Let $\langle k \rangle$ denote the true order statistics for $k = 1, 2, \cdots, K$, such that $g_{\langle 1 \rangle} > g_{\langle 2 \rangle} > \cdots > g_{\langle K \rangle}$.
$\widehat{g}_k^n$	the estimation of g_k at time $0 \le n \le N$. And $\langle k \rangle_n$, for $k = 1, 2, \cdots, K$ and $0 \le n \le N$, denote the estimated order statistics at time n, such that $\widehat{g}_{\langle 1 \rangle_n}^n > \widehat{g}_{\langle 2 \rangle_n}^n > \ldots > \widehat{g}_{\langle K \rangle_n}^n$.

(continued on next page)

(*continued*)

$\boldsymbol{M}_l$	$\boldsymbol{M}_l \triangleq (m_{1,l},\ m_{2,l}, \cdots, m_{K,l})^T$ for $l = 1, \cdots, L$ denote the evaluation result with respect to criterion l. If we choose criterion l to evaluate at time $0 < n \le N$, we represent the result with $\boldsymbol{M}_l^n \triangleq \left(m_{1,l}^n, m_{2,l}^n, \cdots, m_{K,l}^n\right)^T$.
$\boldsymbol{\alpha}_l^n$	$\boldsymbol{\alpha}_l^n \triangleq \left(\alpha_{1,l}^n, \alpha_{2,l}^n, \cdots, \alpha_{K,l}^n\right)^T$, for $l = 1, \cdots, L$ and $0 \le n \le N$, denote the parameter of the posterior distribution of local priority $\boldsymbol{p}_l$ at time n. $\boldsymbol{S}^n \triangleq (\boldsymbol{\alpha}_1^n, \boldsymbol{\alpha}_2^n, \cdots, \boldsymbol{\alpha}_L^n)$ denote the state of knowledge at time $0 \le n \le N$.

From a Bayesian perspective, any number with an unknown value is interpreted as a random variable and begins with a prior distribution updated by the observations (Powell & Ryzhov, 2012). In our problem, experts' evaluation results can be obtained and used to update the distribution of the local priorities.

We make the following assumptions in developing the model.

Assumption 1. Each expert evaluates the designs based on one criterion that they are best at (for example, l) and selects only the best design with respect to criterion l. The selections made by experts satisfy the categorical distribution (a multinomial distribution with the number of trials equal to 1) with parameter $\boldsymbol{p}_l$, i.e., $\boldsymbol{M}_l \sim Cat(\boldsymbol{p}_l)$ for $l = 1, \cdots, L$.

Homogeneity among designs is a prerequisite when compared under the same criterion in AHP (Saaty, 1994; Saaty & Vargas, 2012). Considering the influences of subjective and personal preferences, experts' evaluations may differ. Therefore, we assume that the evaluations satisfied a categorical distribution. In short, when an expert evaluates designs with respect to criterion l, design k is chosen as the best with probability $p_{k,l}$.

Assumption 2. There is an independent Dirichlet prior predictive distribution for local priorities $\boldsymbol{p}_l$, for $l = 1, \cdots, L$, that is, $\boldsymbol{p}_l \sim Dir(\boldsymbol{\alpha}_l^0)$, where $\boldsymbol{\alpha}_l^0 \triangleq \left(\alpha_{1,l}^0, \alpha_{2,l}^0, \cdots, \alpha_{K,l}^0\right)^T$ denotes the parameter of the prior distribution of local priorities $\boldsymbol{p}_l$.

The Dirichlet distribution is a key multivariate distribution for random vectors confined to the simplex and has been widely used in handling data subject to non-negativity and constant-sum constraints (Ng et al., 2011). In our problem, local priorities $\boldsymbol{p}_l$ with respect to criterion l are bounded within the K-dimensional simplex, i.e., $\boldsymbol{p}_l \in T_K(1)$. Herein, $T_K(1) = \left\{ \left(p_{1,l}, p_{2,l}, \cdots, p_{K,l}\right)^T | \sum_{k=1}^K p_{k,l} = 1, p_{k,l} > 0\ \forall 1 \le k \le K \right\}$. Therefore, we assume that local priorities $\boldsymbol{p}_l$ satisfy a Dirichlet prior distribution, which serves as a conjugate prior for the multinomial distribution (Gelman et al., 2014), and we can utilize the experts' evaluation results to update the distribution of local priorities $\boldsymbol{p}_l$.

Assumption 3. The selection of different experts for the same criterion is identically distributed and independent of replication.

Assumption 4. The weights of each criterion, w_l for $l = 1, \cdots, L$, are known in advance.

We assume that the best design is defined as the design with the largest global priority; that is,

$$\langle 1 \rangle \in \underset{k \in \{1,\ldots,K\}}{\operatorname{argmax}} g_k, \tag{1}$$

where $\langle k \rangle$, for $k = 1,2,\cdots,K$, denote the true order statistics. An additive aggregation method is employed to calculate the global priority of each design. The global priority g_k for design k is determined as the weighted average of local priorities $p_{k,l}$, using criterion weights w_l for $l = 1, \cdots, L$, i.e.,

$$g_k = \sum_{l=1}^{L} (w_l \times p_{k,l}),\ \boldsymbol{p}_l = (p_{1,l}, \cdots, p_{K,l})^T \in T_K(1). \tag{2}$$

In practice, local priorities are unknown and estimated by simulation, that is, through experts' evaluations. From a Bayesian perspective, the distribution of local priorities encapsulates our beliefs about the likelihood of local priorities assuming certain values. Therefore, we use the expected values derived from these beliefs to estimate the local priorities:

$$\widehat{p}_{k,l}^n = E^n[p_{k,l}] \text{ and } \widehat{g}_k^n = E^n[g_k] = \sum_{l=1}^{L} (w_l \times E^n[p_{k,l}]). \tag{3}$$

Here, $E^n \triangleq E[\cdot|\boldsymbol{S}^n]$ indicates the conditional expectation with respect to $\boldsymbol{S}^n$, and $\boldsymbol{S}^n = (\boldsymbol{\alpha}_1^n, \boldsymbol{\alpha}_2^n, \cdots, \boldsymbol{\alpha}_L^n)$ denotes the state of knowledge at time $0 \le n \le N$.

2.1. Model formulation

To select the best design using AHP, we calculate the global priorities of the designs, representing their performance, and select the design with the largest global priority. According to Eq. (2), the global priorities are calculated using criterion weights $\boldsymbol{W}$ and local priorities $\boldsymbol{p}_l$ which are unknown in our problem. Therefore, we estimate the local priorities $p_{k,l}$ based on experts' evaluations. With Assumptions 1 and 2, the experts' evaluation results satisfy the categorical distribution with parameter $\boldsymbol{p}_l$, and we use them to update the parameters of the predictive distribution of $\boldsymbol{p}_l$.

Nonetheless, the experts' evaluations may be influenced by individual subjectivity, implying that we must repeat the evaluation with different experts to improve the accuracy of the estimation of local priorities. However, the invitation and evaluation of experts may require considerable time and expense. Thus, we assume that there is a budget N that determines the number of times experts evaluate and choose a criterion for evaluation each time. When the budget is exhausted, we select a design based on the evaluation records and receive an implementation reward.

For each time $0 \le n < N$, we choose a criterion l^n to evaluate at the next time and then observe a corresponding evaluation result $\boldsymbol{M}_{l^n}^{n+1}$, denoting the best design selected by the expert with respect to criterion l^n in time $n+1$, where $\boldsymbol{M}_{l^n}^{n+1} \sim Cat(\boldsymbol{p}_{l^n})$. That is, we obtain a sequence of allocation decisions $l^0, l^1, \cdots, l^n, \cdots, l^{N-1}$ and the corresponding evaluation results $\boldsymbol{M}_{l^0}^1, \boldsymbol{M}_{l^1}^2, \cdots, \boldsymbol{M}_{l^n}^{n+1}, \cdots, \boldsymbol{M}_{l^{N-1}}^N$. Finally, at time N, we determine the design $\langle 1 \rangle_N$ as the "best" design based on the evaluation records.

As our prior distribution is conjugate to the distribution of the evaluation result, the posterior distribution for each $\boldsymbol{p}_l$, where $l = 1, \cdots, L$, is again Dirichlet distributed. That is, at any time $n+1$, we obtain

$$\boldsymbol{p}_l \sim Dir(\boldsymbol{\alpha}_l^n) \text{ for } l = 1, \cdots, L, \text{ and } \boldsymbol{M}_{l^n}^{n+1} | \boldsymbol{p}_{l^n} \sim \mathrm{Cat}(\boldsymbol{p}_{l^n}). \tag{4}$$

According to Powell and Ryzhov (2012), the update function for criterion l^n is given by $\boldsymbol{\alpha}_{l^n}^{n+1} = \boldsymbol{\alpha}_{l^n}^n + \boldsymbol{M}_{l^n}^{n+1}$, and for the criterion that we do not evaluate, our posterior distribution is equal to our prior distribution, i.e., $\boldsymbol{\alpha}_l^{n+1} = \boldsymbol{\alpha}_l^n$ for $l \ne l^n$, that is

$$\boldsymbol{\alpha}_l^{n+1} = \begin{cases} \boldsymbol{\alpha}_l^n + \boldsymbol{M}_l^{n+1}, & \text{if } l = l^n \\ \boldsymbol{\alpha}_l^n, & \text{else} \end{cases} \text{ i.e., } \alpha_{i,l}^{n+1} = \begin{cases} \alpha_{i,l}^n + 1, & \text{if } l = l^n \text{ and } \boldsymbol{M}_{l^n}^{n+1} \text{ belongs to design i} \\ \alpha_{i,l}^n, & \text{else} \end{cases} \tag{5}$$

When our budget is exhausted, the posterior distribution of local priorities $\boldsymbol{p}_l$ is $\mathrm{Dir}(\boldsymbol{\alpha}_l^N)$ for $l = 1, \cdots, L$, and the expectations of the local and global priorities taken over $\boldsymbol{S}^N$ are

$$\widehat{p}_{k,l}^N = \mathrm{E}^N[p_{k,l}] = \frac{\alpha_{k,l}^N}{\sum_{i=1}^K \alpha_{i,l}^N} \text{ and } \widehat{g}_k^N = \mathrm{E}^N[g_k] = \sum_{l=1}^L \left(w_l \times \frac{\alpha_{k,l}^N}{\sum_{i=1}^K \alpha_{i,l}^N} \right). \quad (6)$$

Thus, we opt for the design with the highest estimated global priority, which can be written as

$$\langle 1 \rangle_N = \underset{k\in\{1,..,K\}}{\mathrm{argmax}} \widehat{g}_k^N = \underset{k\in\{1,..,K\}}{\mathrm{argmax}} \mathrm{E}^N[g_k] = \underset{k\in\{1,..,K\}}{\mathrm{argmax}} \sum_{l=1}^L \left(w_l \times \frac{\alpha_{k,l}^N}{\sum_{i=1}^K \alpha_{i,l}^N} \right). \quad (7)$$

Since design $\langle 1 \rangle$ is the actual best design, we use the posterior probability that the AHP correctly selects $\langle 1 \rangle$ as the best design to measure the quality of the evaluation procedure, that is

$$PCS_N \overset{\triangle}{=} P\{\langle 1 \rangle_N = \langle 1 \rangle | \boldsymbol{S}^N\} = P\left\{ \bigcap_{i=2}^K \left(g_{\langle 1 \rangle_N} > g_{\langle i \rangle_N}\right) \middle| \boldsymbol{S}^N \right\}. \quad (8)$$

Defining Z to be a set of all possible allocation policies $\zeta = (l^0, l^1, \cdots, l^{N-1})$, the optimal policy is $\sup_{\zeta \in Z} E[PCS_N|\zeta]$.

2.2. Dynamic allocation

In the evaluation procedure, we choose a criterion to evaluate at each stage that causes a transition in beliefs about local priorities. Thus, the expert allocation problem in the weight valuation step of the AHP is a multistage decision optimization problem that can be solved using a dynamic programming approach.

Using Eq. (8), we measure the evaluation procedure based on the posterior probability that the AHP correctly selects the best design, denoted as PCS_N. Let $V^N(\cdot): \mathrm{S} \rightarrow [0, 1]$ represent the terminal value function for the state $\boldsymbol{S}^N$, that is

$$V^N(\boldsymbol{S}^N) = P\{\langle 1 \rangle_N = \langle 1 \rangle | \boldsymbol{S}^N\} = P\left\{ \bigcap_{i=2}^K \left(g_{\langle 1 \rangle_N} > g_{\langle i \rangle_N}\right) | \boldsymbol{S}^N \right\}. \quad (9)$$

Here, S represents the state space encompassing all possible values of the state variable $\boldsymbol{S}^N$.

At each time $0 \le n < N$, Bellman's equation characterizes the optimal decision using

$$V^n(\boldsymbol{S}^n) = \max_{l\in\{1,\cdots,L\}} E\left[V^{n+1}(\boldsymbol{S}^{n+1}) | \boldsymbol{S}^n, l\right], \quad (10)$$

where $V^n(\cdot)$ is the value function of the state $\boldsymbol{S}^n$, which is composed of $\boldsymbol{\alpha}_l^n$ for $l = 1, \cdots, L$.

According to Eq. (5), we can express state $\boldsymbol{S}^{n+1}$ at time $n+1$ as a function of state $\boldsymbol{S}^n$ at time n and the evaluation result $\boldsymbol{M}_{l^n}^{n+1}$ at time $n+1$. That is, the parameters $\boldsymbol{\alpha}_l^{n+1}$ can be expressed as a function of $\boldsymbol{\alpha}_l^n$ and $\boldsymbol{M}_{l^n}^{n+1}$. At time n, for criteria that are not evaluated at time $n+1$, the parameters $\boldsymbol{\alpha}_l^{n+1}$ remain unchanged. For the criterion selected for evaluation next, the parameter $\boldsymbol{\alpha}_{l^n}^{n+1}$ is a random variable. The predictive distribution of $\boldsymbol{\alpha}_{l^n}^{n+1}$ is discrete, as presented below.

Proposition 1. *At any time $0 \le n < N$, conditionally on $\boldsymbol{S}^n$, if we choose to evaluate criterion l^n at the next iteration $n+1$, the random variable $\boldsymbol{\alpha}_{l^n}^{n+1}$ has a categorical distribution; the probability mass function (PMF) of $\boldsymbol{\alpha}_{l^n}^{n+1}$ is*

$$P\left(\boldsymbol{\alpha}_{l^n}^{n+1} = \boldsymbol{\alpha}_{l^n}^n + \mathbf{e}_k^K | \boldsymbol{S}^n\right) = \frac{\alpha_{k,l^n}^n}{\sum_{i=1}^K \alpha_{i,l^n}^n}, \quad (11)$$

where $\boldsymbol{e}_k^K$ denotes a vector in R^K with all components equal to zero, except for component k, which is equal to 1.

Proof. If we choose to evaluate the criterion l^n, the update function of $\boldsymbol{\alpha}_{l^n}^{n+1}$ is $\boldsymbol{\alpha}_{l^n}^{n+1} = \boldsymbol{\alpha}_{l^n}^n + \boldsymbol{M}_{l^n}^{n+1}$, i.e., $\alpha_{i,l^n}^{n+1} = \alpha_{i,l^n}^n + m_{i,l^n}^{n+1}$.

$$P\left(\boldsymbol{\alpha}_{l^n}^{n+1} = \boldsymbol{\alpha}_{l^n}^n + \boldsymbol{e}_k^K | \boldsymbol{S}^n\right) = P\left(\boldsymbol{M}_{l^n}^{n+1} = \boldsymbol{e}_k^K | \boldsymbol{S}^n\right) = \int_{T_K(1)} P\left(\boldsymbol{M}_{l^n}^{n+1} = \boldsymbol{e}_k^K, \boldsymbol{p}_{l^n} | \boldsymbol{S}^n\right) d\boldsymbol{p}_{l^n}$$

$$= \int_{T_K(1)} P\left(\boldsymbol{M}_{l^n}^{n+1} = \boldsymbol{e}_k^K | \boldsymbol{S}^n, \boldsymbol{p}_{l^n}\right) \times P(\boldsymbol{p}_{l^n} | \boldsymbol{S}^n) d\boldsymbol{p}_{l^n} = \int_{T_K(1)} p_{k,l^n} \times P(\boldsymbol{p}_{l^n} | \boldsymbol{S}^n) d\boldsymbol{p}_{l^n}$$

$$= E\left[p_{k,l^n} | \boldsymbol{S}^n\right] = \frac{\alpha_{k,l^n}^n}{\sum_{i=1}^K \alpha_{i,l^n}^n}.$$

where $\boldsymbol{e}_k^K$ denotes a vector in R^K with all components equal to zero, except for component k, which is equal to 1, and $T_K(1) = \left\{(x_1, x_2, \cdots, x_K)^T | \sum_{k=1}^K x_k = 1, x_k > 0\ \forall 1 \le k \le K\right\}$.■

Because $\boldsymbol{\alpha}_{l^n}^{n+1}$ is a random variable with a categorical distribution under $\boldsymbol{S}^n$, we can express the state in terms of a categorical random variable-adapted sequence $\boldsymbol{M}_{l^0}^1, \boldsymbol{M}_{l^1}^2, \cdots, \boldsymbol{M}_{l^{N-1}}^N$:

$$\boldsymbol{S}^{n+1} = \boldsymbol{S}^n + \boldsymbol{M}_{l^n}^{n+1} \boldsymbol{e}_{l^n}^L, \quad (12)$$

where $\boldsymbol{e}_{l^n}^L$ denotes a vector in R^L with all components equal to zero, except for component l^n, which is equal to 1. We also define a transition function that updates the state by using

$$T(\boldsymbol{S}, l, \boldsymbol{M}_l) = \boldsymbol{S} + \boldsymbol{M}_l \boldsymbol{e}_l^L, \quad (13)$$

where $\boldsymbol{M}_l$ is an evaluation result with respect to criterion l; thus, $\boldsymbol{S}^{n+1} = T(\boldsymbol{S}^n, l^n, \boldsymbol{M}_{l^n}^{n+1})$.

At any time $0 \le n < N$, the optimal decision should be

$$l^n \in \underset{l\in\{1,\cdots,L\}}{\mathrm{argmax}} E\left\{V^{n+1}\left(T\left(\boldsymbol{S}^n, l, \boldsymbol{M}_l^{n+1}\right)\right)\right\}. \quad (14)$$

That is, the criterion l^n that maximizes the expectation value of $V^{n+1}\left(T\left(\boldsymbol{S}^n, l, \boldsymbol{M}_l^{n+1}\right)\right)$ is evaluated at time $n+1$.

3. The KG policy of AHP

Solving for the optimal decision in dynamic programming is virtually impossible because of the complexity of Eq. (14). In this section, we generalize the concept of the KG policy, as introduced by Frazier et al. (2008), to address the expert allocation problem in the AHP, namely, the AHP-KG policy.

3.1. Derivation

As Frazier et al. (2008) state, the KG policy formulates an equivalent problem by dividing the reward into each period of the procedure, ensuring that the total reward received is identical to that of the original problem, that is,

$$V^N(\boldsymbol{S}^N) = \left[V^N(\boldsymbol{S}^N) - V^N(\boldsymbol{S}^{N-1})\right] + \cdots + \left[V^N(\boldsymbol{S}^{n+1}) - V^N(\boldsymbol{S}^n)\right] + V^N(\boldsymbol{S}^n). \quad (15)$$

The KG policy provides a one-step Bayes-optimal point for evaluation based on the assumption that the evaluation procedure will be terminated immediately afterwards (Groves & Branke, 2019). At each time, the KG policy chooses the criterion to evaluate the next, which maximizes the expected increase in a single-period in the terminal value function, i.e., $\mathrm{argmax}_{l\in\{1,\cdots,L\}} E^n\left[V^N\left(\boldsymbol{S}^{n+1}\right) - V^N(\boldsymbol{S}^n) | l^n = l\right]$.

At any time $0 \le n < N$, the AHP-KG policy is defined to choose the

Algorithm 1
AHP-KG policy.

INPUT: The total simulation budget N.
The parameters of the prior Dirichlet distribution $\boldsymbol{\alpha}_l$ for $l = 1, \cdots, L$.
The weights of criteria w_l for $l = 1, \cdots, L$.
INITIALIZE: The iteration counter $n \leftarrow 0$.
The state of knowledge $\boldsymbol{S} = (\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_L)$.
WHILE $n < N$ **DO**:
Calculate the approximated AHP-KG factor $\widetilde{v}_l^{KG}(\boldsymbol{S})$ of each criterion l with Eq. (20).
$l^n \leftarrow \mathrm{argmax}_{l \in \{1, \cdots, L\}} \widetilde{v}_l^{KG}(\boldsymbol{S})$.
Invite an additional expert to evaluate the designs with respect to criterion l^n, and obtain the result $\boldsymbol{M}_{l^n}$ based on this expert's evaluation.
$\boldsymbol{\alpha}_{l^n} \leftarrow \boldsymbol{\alpha}_{l^n} + \boldsymbol{M}_{l^n}$.
$n \leftarrow n + 1$.
END WHILE
Select the design with the largest estimated global priority as the final decision.

Algorithm 2
AHP-AKG policy.

INPUT: The total simulation budget N.
The parameters of the prior Dirichlet distribution $\boldsymbol{\alpha}_l$ for $l = 1, \cdots, L$.
INITIALIZE: The iteration counter $n \leftarrow 0$.
The number of evaluation results $r \leftarrow 0$.
The number of workers currently running an evaluation for each criterion $h_l \leftarrow 0$ for $l = 1, 2, \cdots, L$.
The status of each worker *worker.status*←*Free* for all *worker* ∈C.
The state of knowledge $\boldsymbol{S} = (\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_L)$.
IF *worker.status* equals *Free* for any *worker c* **THEN**:
Select a criterion randomly (e.g., criterion d), and invite an expert to evaluate with respect to criterion d.
worker.status←*Working* for *worker c*.
$h_d \leftarrow h_d + 1$.
$n \leftarrow n + 1$.
END IF
WHILE $n < N$ **DO**:
IF *worker.status* equals *Completed* for any *worker c* **THEN**:
$r \leftarrow r + 1$.
Obtain the $r-$th result $\boldsymbol{M}_d$ from *worker c*.
$h_d \leftarrow h_d - 1$.
Update the parameters of Dirichlet distribution $\boldsymbol{\alpha}_d = \boldsymbol{\alpha}_d + \boldsymbol{M}_d$ for criterion d.
Calculate the value of $\min\left[\left(c_{2,l}^{n+1}\right)^2, \cdots, \left(c_{K,l}^{n+1}\right)^2\right]$ for each criterion l with Eq. (24).
Find criterion $l^n \in \mathrm{argmax}_{l \in \{1, \cdots, L\}} \left\{\min\left[\left(c_{2,l}^{n+1}\right)^2, \cdots, \left(c_{K,l}^{n+1}\right)^2\right]\right\}$.
Let *worker c* invite an additional expert to evaluate the designs with respect to criterion l^n.
worker.status←*Working* for *worker c*.
$h_{l^n} \leftarrow h_{l^n} + 1$.
$n \leftarrow n + 1$.
END IF
END WHILE
Select the design with the largest estimated global priority as the final decision.

Table 1
Local priorities for the selection problem of rough terrain cargo handlers.

designs	Performance *Weight* 0.517	Risk 0.059	Cost 0.306	Program Objective 0.118
Baseline	0.187	0.640	0.490	0.278
Upgraded System	0.229	0.283	0.335	0.259
USDCH	0.584	0.077	0.175	0.463

criterion $l^{KG}(\boldsymbol{S}^n)$ that maximizes the expectation of a single-period reward with respect to $\boldsymbol{S}^n$, i.e.,

$$l^{KG}(\boldsymbol{S}^n) \in \underset{l \in \{1, \cdots, L\}}{\mathrm{argmax}} E^n\left[V^N\left(T\left(\boldsymbol{S}^n, l, \boldsymbol{M}_l^{n+1}\right)\right) - V^N(\boldsymbol{S}^n)\right]. \tag{16}$$

As the value of $E^n\left[V^N(\boldsymbol{S}^n)\right]$ is independent of l^n, the allocation rule for the AHP-KG policy can be expressed as

$$l^{KG}(\boldsymbol{S}^n) \in \underset{l \in \{1, \cdots, L\}}{\mathrm{argmax}} E^n\left[V^N\left(T\left(\boldsymbol{S}^n, l, \boldsymbol{M}_l^{n+1}\right)\right)\right]. \tag{17}$$

We define the AHP-KG factor $v_l^{KG}(\boldsymbol{S}^n)$ for criterion l at state $\boldsymbol{S}^n$ as the expected terminal value of the subsequent state when evaluating criterion l next. The AHP-KG policy chooses the criterion that maximizes the AHP-KG factor to be evaluated next:

$$v_l^{KG}(\boldsymbol{S}^n) = E^n\left[V^N\left(T\left(\boldsymbol{S}^n, l, \boldsymbol{M}_l^{n+1}\right)\right)\right], \tag{18}$$

$$l^{KG}(\boldsymbol{S}^n) \in \underset{l \in \{1, \cdots, L\}}{\mathrm{argmax}} v_l^{KG}(\boldsymbol{S}^n). \tag{19}$$

The following theorem provides an efficient method for approximating the AHP-KG factor $v_l^{KG}(\boldsymbol{S}^n)$ and the associated AHP-KG policy.

Theorem 1. *At any time $0 \le n < N$, conditionally on $\boldsymbol{S}^n$, the AHP-KG factor for criterion l can be approximated by*

$$\widetilde{v}_l^{KG}(\boldsymbol{S}^n) \triangleq \frac{\pi \times \min\left[\left(d_{2,l}^{n+1}\right)^2, \cdots, \left(d_{K,l}^{n+1}\right)^2\right]}{2}, \tag{20}$$

where

$$\left(d_{k,l}^{n+1}\right)^2 \triangleq \frac{\left[\sum_{d=1}^{L} w_d \times \left(\frac{\alpha_{\langle 1\rangle_n,d}^{n} - \alpha_{\langle k\rangle_n,d}^{n}}{\sum_{i=1}^{K}\alpha_{i,d}^{n}}\right)\right]^2}{\sum_{d=1}^{L}\left[w_d^2 \times \frac{\tilde{\alpha}_{\langle 1\rangle_n,d}^{n+1} \times \left(\sum_{i=1}^{K}\tilde{\alpha}_{i,d}^{n+1} - \tilde{\alpha}_{\langle 1\rangle_n,d}^{n+1}\right) + \tilde{\alpha}_{\langle k\rangle_n,d}^{n+1} \times \left(\sum_{i=1}^{K}\tilde{\alpha}_{i,d}^{n+1} - \tilde{\alpha}_{\langle k\rangle_n,d}^{n+1}\right) + 2 \times \tilde{\alpha}_{\langle 1\rangle_n,d}^{n+1} \times \tilde{\alpha}_{\langle k\rangle_n,d}^{n+1}}{\left(\sum_{i=1}^{K}\tilde{\alpha}_{i,d}^{n+1}\right)^2 \times \left(\sum_{i=1}^{K}\tilde{\alpha}_{i,d}^{n+1} + 1\right)}\right]},$$

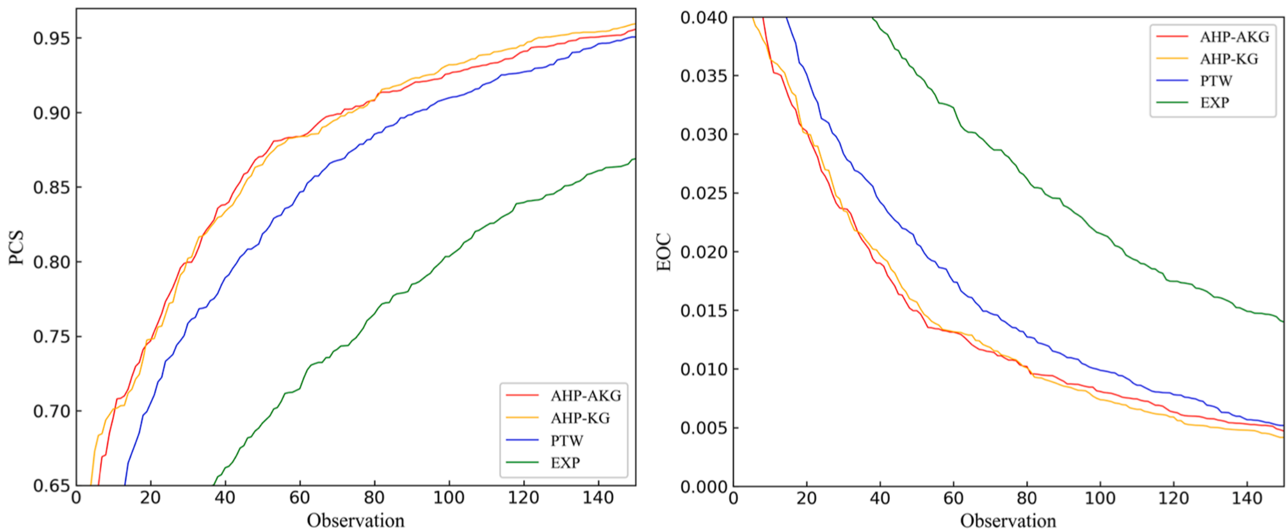

Fig. 2. The PCS and EOC for the selection problem with Dirichlet distribution.

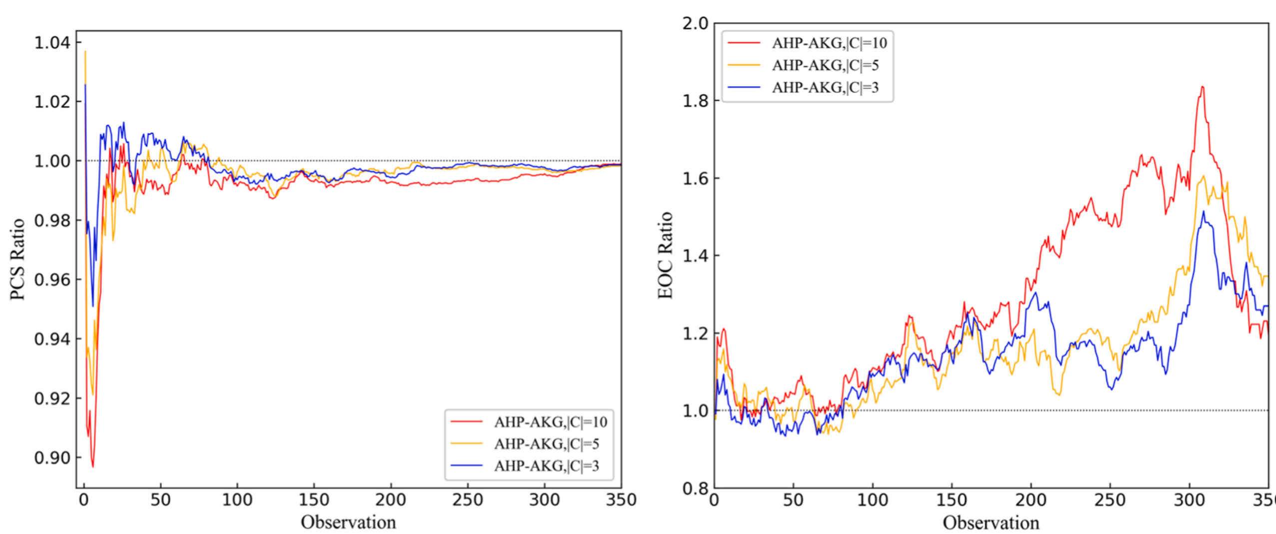

Fig. 3. The PCS and EOC ratios for the AHP-AKG policy with different numbers of workers denoted by $|C|$. On the left penal, it displays the ratio of the PCS of the AHP-AKG policy to that of the AHP-KG policy, and the right panel is the ratio of the EOC.

Table 2
Parameters setting for different experiments.

Scenarios	L	K	w_l	$p_{k,l}$	α_l^0
1	10	3	$(11-l)/55$	$(k+l-1)/\sum_k(k+l-1)$	$\mathbf{1}_K$
2	10	3	$(11-l)/55$	$(k+l-1)/\sum_k(k+l-1)$	$(10,9,8)^T$
3	10	3	$[10-5(-1)^l]/100$	$[2k+(-1)^l]/\sum_k[2k+(-1)^l]$	$\mathbf{1}_K$
4	10	3	$(11-l)/55$	$[2l+(-1)^k]/\sum_k[2l+(-1)^k]$	$\mathbf{1}_K$
5	20	3	$(21-l)/210$	$(k+l-1)/\sum_k(k+l-1)$	$\mathbf{1}_K$
6	10	5	$(11-l)/55$	$(k+l-1)/\sum_k(k+l-1)$	$\mathbf{1}_K$

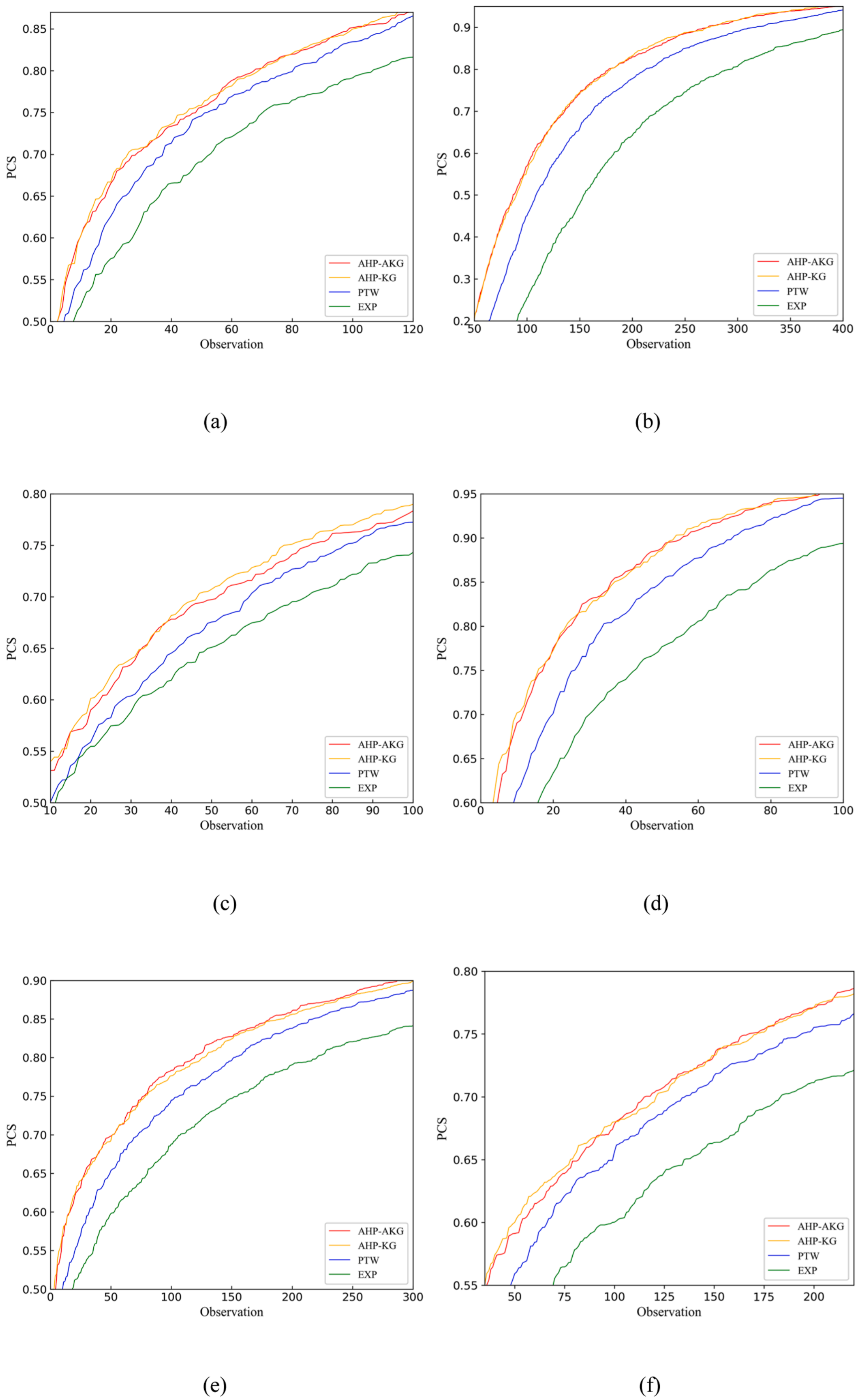

Fig. 4. Performance comparison of PCS under different scenarios. Subplots (a)–(f) are the results of scenarios 1–6.

$$\widetilde{\boldsymbol{\alpha}}_l^{n+1} \triangleq \left(\widetilde{\alpha}_{1,l}^{n+1}, \cdots, \widetilde{\alpha}_{K,l}^{n+1}\right)^T = \left(\alpha_{1,l}^n \times \frac{\sum_{i=1}^K \alpha_{i,l}^n + 1}{\sum_{i=1}^K \alpha_{i,l}^n}, \cdots, \alpha_{K,l}^n \times \frac{\sum_{i=1}^K \alpha_{i,l}^n + 1}{\sum_{i=1}^K \alpha_{i,l}^n}\right)^T,$$

$\widetilde{\boldsymbol{\alpha}}_i^{n+1} \triangleq \left(\widetilde{\alpha}_{1,i}^{n+1}, \cdots, \widetilde{\alpha}_{K,i}^{n+1}\right)^T = \boldsymbol{\alpha}_i^n$ for $i \neq l$, and π represents Pi.

The AHP-KG allocation rule is

$$l^{KG}(\boldsymbol{S}^n) \in \underset{l \in \{1,\cdots,L\}}{\operatorname{argmax}} \left\{ \min\left[\left(d_{2,l}^{n+1}\right)^2, \cdots, \left(d_{K,l}^{n+1}\right)^2 \right] \right\}. \tag{21}$$

Proof. See Appendix A.

3.2. Sequential evaluation procedure

Using Theorem 1, we present a sequential procedure based on the AHP-KG policy to select the best design from K designs, given a specified evaluation budget. Our problem is parameterized by the initial parameters of the Dirichlet distribution $\boldsymbol{\alpha}_l^0$ for $l = 1, \cdots, L$, the weights of criteria w_l for $l = 1, \cdots, L$, and the total evaluation budget N. For each iteration, we allocate an expert to the criterion with the largest approximated AHP-KG factor $\widetilde{v}_l^{KG}(\boldsymbol{S}^n)$ and update the parameters of the predictive Dirichlet distribution until the budget is exhausted. Finally, we select the design with the highest estimated global priority. In summary, we present Algorithm 1.

3.3. Asymptotic optimality

A desirable property of the expert allocation policy in AHP is *Asymptotic Optimality*, which ensures convergence to the best design, given an infinite evaluation budget (Groves & Branke, 2019). For instance, the pure exploration policy samples a criterion with a probability $1/L$, and the proportional to weights allocation policy allocates experts based on the criterion weights $\boldsymbol{W} = (w_1, w_2, \cdots, w_L)^T$. These policies are asymptotically optimal. Given the purely myopic nature of the AHP-KG policy, proving its asymptotic optimality is crucial.

The asymptotic optimality of the KG policy for the standard ranking and selection problem is well-established (Frazier & Powell, 2008), where simulations directly estimate the performance of designs with unbounded sample responses of finite variance. In Theorem 2, we demonstrate that the asymptotic optimality is sustained for the AHP-KG policy.

Theorem 2. *The knowledge gradient policy for the expert allocation problem in AHP (AHP-KG policy) is asymptotically optimal.*

Proof. See Appendix B.

4. The AKG policy of AHP

4.1. Derivation

For the expert allocation problem in the weight valuation step of using AHP to select the best design, inviting experts and conducting evaluations typically requires extended periods of time and can vary in duration. There is also a risk of failure. The AKG policy introduced by Kaminski and Szufel (2014) is an asynchronous parallelized version of the KG policy. The asynchronous policy assumes that simulations or evaluations are executed on several heterogeneous workers and is designed to be resilient to failures among workers.

In the expert allocation problem, we assume that $|\mathrm{C}|$ workers complete the tasks of inviting experts and collecting evaluation results simultaneously, where C denotes the pool of workers. When a worker finishes its task, it is immediately assigned a new task based on the previously collected evaluation results and information about tasks that were started but had not yet been finalized.

At any time $0 \le n < N$, we need to choose a criterion l^n from the set $\{1, 2, \cdots, L\}$ to be evaluated at the next time. However, it should be noted that workers may still complete evaluations initiated earlier but not yet finalized. Suppose we obtain the r-th evaluation result, the worker that provided it is waiting for the next job, and the evaluation results from $r+1$ to n are unknown. We denote the number of workers currently evaluating criterion l as h_l^n, that is, $n = r + \sum_{l=1}^L h_l^n$. Here, the value $\boldsymbol{H}^n \triangleq (h_1^n, \cdots, h_L^n)$ is known.

At any time $0 \le n < N$, we only obtain r evaluation results that construct the state $\boldsymbol{S}^r$. Following the AHP-KG policy approach, we define the AHP-AKG factor for criterion l as the expectation of the terminal value of the next state with respect to $\boldsymbol{S}^r$ and $\boldsymbol{H}^n$, i.e.,

$$v_l^{AKG}(\boldsymbol{S}^r, \boldsymbol{H}^n) \triangleq E^r\left[V^N\left(T\left(\boldsymbol{S}^n, l, \boldsymbol{M}_l^{n+1}\right)\right) \middle| \boldsymbol{H}^n\right]. \tag{22}$$

We choose the criterion that maximizes the value of $v_l^{AKG}(\boldsymbol{S}^r, \boldsymbol{H}^n)$, that is

$$l^{AKG}(\boldsymbol{S}^r, \boldsymbol{H}^n) \in \underset{l \in \{1,\cdots,L\}}{\operatorname{argmax}} E^r\left[V^N\left(T\left(\boldsymbol{S}^n, l, \boldsymbol{M}_l^{n+1}\right)\right) \middle| \boldsymbol{H}^n\right]. \tag{23}$$

The allocation rule of the AHP-AKG policy is shown below.

Theorem 3. *At any time $0 \le n < N$, conditionally on $\boldsymbol{S}^r$ and $\boldsymbol{H}^n$, the AHP-AKG allocation rule is*

$$l^{AKG}(\boldsymbol{S}^r, \boldsymbol{H}^n) \in \underset{l \in \{1,\cdots,L\}}{\operatorname{argmax}} \left\{ \min\left[\left(c_{2,l}^{n+1}\right)^2, \cdots, \left(c_{K,l}^{n+1}\right)^2 \right] \right\}, \tag{24}$$

where

$$\left(c_{k,l}^{n+1}\right)^2 \triangleq \frac{\left[\sum_{d=1}^L w_d \times \left(\frac{\alpha_{\langle 1\rangle_r,d}^r - \alpha_{\langle k\rangle_r,d}^r}{\sum_{i=1}^K \alpha_{i,d}^r}\right)\right]^2}{\sum_{d=1}^L \left[w_d^2 \times \frac{\widehat{\alpha}_{\langle 1\rangle_r,d}^{n+1} \times \left(\sum_{i=1}^K \widehat{\alpha}_{i,d}^{n+1} - \widehat{\alpha}_{\langle 1\rangle_r,d}^{n+1}\right) + \widehat{\alpha}_{\langle k\rangle_r,d}^{n+1} \times \left(\sum_{i=1}^K \widehat{\alpha}_{i,d}^{n+1} - \widehat{\alpha}_{\langle k\rangle_r,d}^{n+1}\right) + 2 \times \widehat{\alpha}_{\langle 1\rangle_r,d}^{n+1} \times \widehat{\alpha}_{\langle k\rangle_r,d}^{n+1}}{\left(\sum_{i=1}^K \widehat{\alpha}_{i,d}^{n+1}\right)^2 \times \left(\sum_{i=1}^K \widehat{\alpha}_{i,d}^{n+1} + 1\right)} \right]},$$

and $\widehat{\boldsymbol{\alpha}}_l^{n+1} \triangleq \left(\widehat{\alpha}_{1,l}^{n+1}, \cdots, \widehat{\alpha}_{K,l}^{n+1}\right)^T = \left(\alpha_{1,l}^r \times \frac{\sum_{i=1}^K \alpha_{i,l}^r + h_l^n + 1}{\sum_{i=1}^K \alpha_{i,l}^r}, \cdots, \alpha_{K,l}^r \times \frac{\sum_{i=1}^K \alpha_{i,l}^r + h_l^n + 1}{\sum_{i=1}^K \alpha_{i,l}^r}\right)^T$,
$\widehat{\boldsymbol{\alpha}}_j^{n+1} \triangleq \left(\widehat{\alpha}_{1,j}^{n+1}, \cdots, \widehat{\alpha}_{K,j}^{n+1}\right)^T = \left(\alpha_{1,j}^r \times \frac{\sum_{i=1}^K \alpha_{i,j}^r + h_j^n}{\sum_{i=1}^K \alpha_{i,j}^r}, \cdots, \alpha_{K,j}^r \times \frac{\sum_{i=1}^K \alpha_{i,j}^r + h_j^n}{\sum_{i=1}^K \alpha_{i,j}^r}\right)$ for $j \neq l$.

Proof. The proof of Theorem 3 is similar to that of Theorem 1. To avoid repetition, we omit the proof.

4.2. Sequential evaluation procedure

According to Theorem 3, we can develop a sequential procedure for weight valuation using AHP to select the best design based on the AHP-AKG policy.

In the evaluation procedure, we assume the presence of one master machine controlling the entire evaluation process and a set of $|\mathrm{C}|$ worker machines assigned to the tasks of inviting experts and collecting evaluation results. A worker begins to invite an expert upon receiving instructions from the master. Once the worker obtains the evaluation result, it reports to the master. The master then updates the parameters of the predictive Dirichlet distribution, approximates the AHP-AKG factors as presented in Theorem 3, and subsequently reassigns a new task to the worker.

In practice, a single evaluation task may fail or a worker might break down. If a worker fails to execute an evaluation task, the master assigns a new task to the worker according to the AHP-AKG allocation rule. When a worker breaks down, it is removed from C. Similarly, new workers can be added to C at any time. For simplicity, we assume that no new workers are added and no worker breaks down. Each *worker* within C possesses one of the following three statuses:

(a) *Free* – a new created worker.
(b) *Working* – an invitation or evaluation of an expert is in progress.
(c) *Completed* – an invitation or evaluation of an expert is complete and can take up a new task.

The full procedure of the AHP-AKG policy is presented in Algorithm 2.

5. Numerical experiments

To evaluate the efficiency improvement by the AHP-KG and AHP-AKG policies, we compare them with the pure exploration and proportional to weight allocation policies, denoted as the EXP and PTW policies, respectively. At each step, the EXP policy samples the criteria with probability $1/L$, uses the result of the expert evaluation to update the parameters of the Dirichlet distributions according to Eq. (5), and selects the best design based on Eq. (3). In the PTW policy, the experts are allocated based on the weights of the criterion $\boldsymbol{W} = (w_1, w_2, \cdots, w_L)^T$.

In all numerical experiments, we independently repeat the procedure 10,000 times to obtain the results. We estimate two metrics, the probability of correct selection (PCS) and the expected opportunity cost (EOC), to measure the quality of the selection in the evaluation procedure. The PCS is calculated as the ratio of the successful identification of the true best design to the total number of procedures, representing the frequency of correct selection. The EOC represents the mean difference between the selection of the procedure and optimal design, indicating the extent to which the selection of the evaluation procedure may be suboptimal.

5.1. Experiment based on the real case study

To test the effectiveness of the AHP-KG and AHP-AKG policies, we apply them to the real case of rough terrain cargo handler selection problem, as demonstrated in the literature (Bard & Sousk, 1990; Huang et al., 2015) and shown in Fig. 1. Contrary to our assumption, the selection problem in Huang et al. (2015) assumes that the local priorities deduced by the PCMs are normally distributed. Considering the background of our study, where experts only select the best design with respect to one criterion, and local priorities are estimated without PCMs, we treat the mean of the normal distributions as the local priorities, as shown in Table 1.

The prior parameters of the local priorities $\boldsymbol{\alpha}_l^0$ for $l = 1, \cdots, L$ are set to $\mathbf{1}_K$, which means a K-dimensional vector of ones to represent the state with no information about the local priorities. At each iteration, according to Theorems 2 and 3, we calculate the AHP-KG and AHP-AKG factors and select the criterion with the largest factor to be evaluated in the next iteration. If we decide to evaluate the designs with respect to criterion *l*, we sample the evaluation result from the multinomial distribution with parameters equal to the local priorities for criterion *l*, as listed in Table 1. Subsequently, we use the evaluation result to update the parameters of the Dirichlet distribution for criterion *l* according to Eq. (5), and the AHP-KG and AHP-AKG factors are recalculated based on the updated parameters. This procedure is repeated until the evaluation budget is exhausted, and the design with the largest estimated global priority is selected as the best design.

As shown in Fig. 2, the number of observations on the horizontal axis represents the number of evaluation results obtained during the procedure. For the policies, each observation provides information on the best design with respect to one criterion. PCS increases with the number of observations in the experiments, whereas EOC decreases. In particular, the performances of the AHP-KG and AHP-AKG policies are the best in this experiment. When determining the criterion to be evaluated in the next iteration, the AHP-KG and AHP-AKG policies exploit information regarding both the criterion weights and the means and variances of the posterior distribution. The means and variances provide valuable information regarding the relative differences across the design space.

In Fig. 2, the performance difference between the AHP-AKG policy with three workers and the AHP-KG policy (which is the AHP-AKG policy with one worker) is very small. To evaluate how varying numbers of workers affect performance, in the left panel of Fig. 3, we plot the ratio of the PCS of the AHP-AKG policy to that of the AHP-KG policy. The right panel compares the ratio of EOC. The difference in the PCS for different numbers of workers decreases as the number of evaluations increases, and the opposite is true for the EOC.

5.2. Testing on different scenarios

We will explore the relative quality of the AHP-KG policy, the AHP-AKG policy with 3 workers, and the PTW, EXP policies using numerical simulations across several test cases. The detailed parameter settings are listed in Table 2.

In Table 2, L and K represent the numbers of criteria and designs, respectively. We conduct experiments with 10 or 20 criteria, that is, $L \in \{10, 20\}$, and the number of designs is 3 or 5, that is, $K \in \{3, 5\}$. The w_l for $l = 1, 2, \cdots, L$ is the value of the criterion weight, and $p_{k,l}$ for $l = 1, 2, \cdots, L$ and $k = 1, 2, \cdots, K$ denotes the local priority of design k with respect to criterion l. The prior parameters of the local priorities $\boldsymbol{\alpha}_l^0$ for $l = 1, \cdots, L$ are set to $\mathbf{1}_K$, except for scenario 2, where the prior parameters $\boldsymbol{\alpha}_l^0$ for $l = 1, \cdots, L$ are equal to $(10, 9, 8)^T$ to test the performance of our policies when initial priors provide misleading information about the local priorities. For scenarios 3 and 4, we consider situations with significant differences in criterion weights and local priorities. In scenarios 5 and 6, different numbers of criteria and designs are considered.

Fig. 4 shows that the AHP-KG and AHP-AKG policies outperform the PTW and EXP policies across all scenarios. Notably, in scenario 2, which is characterized by inaccurate initial local priority estimates, the advantages of the AHP-KG and AHP-AKG policies remain obvious during the evaluation procedures.

According to the results from these scenarios, the AHP-AKG and AHP-KG policies are the best performers among the compared methods, and the difference in performance between them is insignificant. The AHP-AKG policy's parallelization capability allows for the simultaneous evaluation by three experts, which makes the evaluation process three times faster. Given that the performance of the AHP-AKG policy is no worse than that of the AHP-KG policy, we conclude that the benefits of parallelization are significant. Therefore, with multiple workers available, the parallelization and asynchrony of the AHP-AKG policy can lead to significant time savings.

6. Conclusion and future work

In this study, we address the problem of using AHP to select the best design from a finite number of K designs. The local priorities are estimated from a Bayesian perspective using the Dirichlet-multinomial model. In our problem, experts only need to select the best design based on their expertise, and the parameters of the local priority distributions are updated according to the Bayesian rule. This improves the accuracy and efficiency of experts' evaluations, allowing us to estimate local priorities without PCMs.

Building on this foundation, we address the expert allocation problem during the weight valuation step of the AHP. We propose two expert allocation policies, AHP-KG and AHP-AKG, to improve the efficiency of the evaluation procedure. Our numerical results show that the AHP-KG and AHP-AKG policies can improve the PCS within a specified evaluation budget. Furthermore, the AHP-AKG policy, with its parallelization and asynchrony, can significantly improve the efficiency of the weight valuation step in the AHP.

A possible future development would be to consider the probability of good selection (PGS) of the AHP procedure, especially when decision-makers find a suboptimal design acceptable. Additionally, the concept of expert allocation policy can be extended to other performance metrics such as computational time, especially when dealing with a large number of designs. In this context, a significant concern is the computational costs associated with allocation rules, which impose limitations on Bayesian selection procedures (Eckman & Henderson, 2022).

Another promising direction for future research involves exploring alternative estimation methods, such as interval estimation and robust estimation. The robust estimation method described by Gao et al. (2017) considers input uncertainty. In such cases, the worst-case design performance is used to establish a robust selection rule.

CRediT authorship contribution statement

Hui Xiao: Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Conceptualization. **Sha Zeng:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation. **Yi Peng:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Funding acquisition, Formal analysis, Conceptualization. **Gang Kou:** Writing – review & editing, Writing – original draft, Supervision, Funding acquisition, Conceptualization.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China under grant numbers 72431008 and 72371201, and the Science and Technology Innovation Program of Hunan Province under grant number 2024RC4008.

Appendix A: Proof of Theorem 1

Proof. At any time $0 \le n < N$, the posterior distribution of the local priorities $\boldsymbol{p}_l = \left(p_{1,l}, p_{2,l}, \cdots, p_{K,l}\right)^T$ is $Dir(\boldsymbol{\alpha}_l^n)$ for $l = 1, \cdots, L$. According to Ng et al. (2011), the marginal distribution of $p_{k,l}$ is a beta distribution, that is, $p_{k,l} \sim Beta\left(\alpha_{k,l}^n, \sum_{i=1}^K \alpha_{i,l}^n - \alpha_{k,l}^n\right)$. The covariance between $p_{k,l}$ and $p_{j,l}$ is $Cov\left(p_{k,l}, p_{j,l}\right) = -\frac{\alpha_{k,l}^n \times \alpha_{j,l}^n}{\left(\sum_{i=1}^K \alpha_{i,l}^n\right)^2 \times \left(\sum_{i=1}^K \alpha_{i,l}^n + 1\right)}$. As shown in Li et al. (2019), the beta distribution $Beta(\alpha, \beta)$ converges in distribution to the normal distribution $N\left(\frac{\alpha}{(\alpha+\beta)}, \frac{\alpha \times \beta}{(\alpha+\beta)^2 \times (\alpha+\beta+1)}\right)$. Since beta distributions are not closed in a linear combination, we use the normal distribution to approximate the posterior distribution of $p_{k,l}$, that is

$$p_{k,l} \sim N\left(\frac{\alpha_{k,l}^n}{\sum_{i=1}^K \alpha_{i,l}^n}, \frac{\alpha_{k,l}^n \times \left(\sum_{i=1}^K \alpha_{i,l}^n - \alpha_{k,l}^n\right)}{\left(\sum_{i=1}^K \alpha_{i,l}^n\right)^2 \times \left(\sum_{i=1}^K \alpha_{i,l}^n + 1\right)}\right).$$

We use the additive aggregation method to calculate the global priority of each design, that is, $g_k = \sum_{l=1}^L (w_l \times p_{k,l})$, and following Assumption 2, the local priorities across the different criteria are independent. Subsequently, the posterior approximation distribution of g_k, which is a linear combination of normal distributions, follows a normal distribution:

$$g_k \sim N\left(\sum_{l=1}^L \left(w_l \times \frac{\alpha_{k,l}^n}{\sum_{i=1}^K \alpha_{i,l}^n}\right), \sum_{l=1}^L \left[w_l^2 \times \frac{\alpha_{k,l}^n \times \left(\sum_{i=1}^K \alpha_{i,l}^n - \alpha_{k,l}^n\right)}{\left(\sum_{i=1}^K \alpha_{i,l}^n\right)^2 \times \left(\sum_{i=1}^K \alpha_{i,l}^n + 1\right)}\right]\right).$$

The covariance between g_k and g_j for $j \neq k$ is

$$Cov\left(g_k, g_j\right) = Cov\left(\sum_{l=1}^L (w_l \times p_{k,l}), \sum_{l=1}^L \left(w_l \times p_{j,l}\right)\right) = \sum_{l=1}^L \left[w_l \times Cov\left(p_{k,l}, \sum_{l=1}^L \left(w_l \times p_{j,l}\right)\right)\right]$$

$$= \sum_{l=1}^L \left[w_l^2 \times Cov\left(p_{k,l}, p_{j,l}\right)\right] = -\sum_{l=1}^L \left[w_l^2 \times \frac{\alpha_{k,l}^n \times \alpha_{j,l}^n}{\left(\sum_{i=1}^K \alpha_{i,l}^n\right)^2 \times \left(\sum_{i=1}^K \alpha_{i,l}^n + 1\right)}\right].$$

According to Eq. (9), we have

$$V^N(\boldsymbol{S}^n) = P\{\langle 1\rangle_n = \langle 1\rangle | \boldsymbol{S}^n\} = P\left\{\bigcap_{i=2}^K \left(g_{\langle 1\rangle_n} > g_{\langle i\rangle_n}\right) | \boldsymbol{S}^n\right\} = P\left\{\bigcap_{i=2}^K \left(g_{\langle 1\rangle_n} - g_{\langle i\rangle_n} > 0\right) | \boldsymbol{S}^n\right\}.$$

The joint distribution of vector $\left(g_{\langle 1\rangle_n} - g_{\langle 2\rangle_n}, \cdots, g_{\langle 1\rangle_n} - g_{\langle K\rangle_n}\right)$ follows a joint normal distribution with mean vector $\left(\sum_{l=1}^{L}\left(w_l \times \frac{\alpha^n_{\langle 1\rangle_n,l} - \alpha^n_{\langle 2\rangle_n,l}}{\sum_{i=1}^{K}\alpha^n_{i,l}}\right), \cdots, \sum_{l=1}^{L}\left(w_l \times \frac{\alpha^n_{\langle 1\rangle_n,l} - \alpha^n_{\langle K\rangle_n,l}}{\sum_{i=1}^{K}\alpha^n_{i,l}}\right)\right)$. The covariance matrix is given by

$$\Xi \triangleq \begin{pmatrix} T_2 & R_{2,3} & \cdots & R_{2,K} \\ R_{3,2} & T_3 & \cdots & R_{3,K} \\ \vdots & \vdots & & \vdots \\ R_{K,2} & R_{K,3} & \cdots & T_K \end{pmatrix}_{(K-1)\times(K-1)},$$

where $T_k = \sum_{l=1}^{L}\left[w_l^2 \times \frac{\alpha^n_{\langle 1\rangle_n,l}\times\left(\sum_{i=1}^{K}\alpha^n_{i,l} - \alpha^n_{\langle 1\rangle_n,l}\right) + \alpha^n_{\langle k\rangle_n,l}\times\left(\sum_{i=1}^{K}\alpha^n_{i,l} - \alpha^n_{\langle k\rangle_n,l}\right) + 2\times\alpha^n_{\langle 1\rangle_n,l}\times\alpha^n_{\langle k\rangle_n,l}}{\left(\sum_{i=1}^{K}\alpha^n_{i,l}\right)^2\times\left(\sum_{i=1}^{K}\alpha^n_{i,l}+1\right)}\right]$, and $R_{i,j} = R_{j,i} = \sum_{l=1}^{L}\left[w_l^2 \times \frac{\alpha^n_{\langle 1\rangle_n,l}\times\left(\sum_{i=1}^{K}\alpha^n_{i,l} - \alpha^n_{\langle 1\rangle_n,l}\right) + \alpha^n_{\langle 1\rangle_n,l}\times\alpha^n_{\langle i\rangle_n,l} + \alpha^n_{\langle 1\rangle_n,l}\times\alpha^n_{\langle j\rangle_n,l} - \alpha^n_{\langle i\rangle_n,l}\times\alpha^n_{\langle j\rangle_n,l}}{\left(\sum_{i=1}^{K}\alpha^n_{i,l}\right)^2\times\left(\sum_{i=1}^{K}\alpha^n_{i,l}+1\right)}\right]$.

The covariance matrix Ξ is a semi-positive matrix, and we regularize it by $\lambda\cdot\mathrm{I}$, where λ is a small positive number and I is the identity matrix, i.e., $\Gamma \triangleq \Xi + \lambda\cdot\mathrm{I}$. By Cholesky decomposition, we have $\Gamma = U'U$, where $U = [u_{i,j}]_{(K-1)\times(K-1)}$ is an upper triangular matrix. Thus, the value of $V^N(\boldsymbol{S}^n)$ can be approximated as:

$$V^N(\boldsymbol{S}^n) = P\left\{\bigcap_{i=2}^{K} g_{\langle 1\rangle_n} - g_{\langle i\rangle_n} > 0 \middle| \boldsymbol{S}^n\right\} \approx P\left\{\sum_{j=1}^{i-1} u_{j,i-1}Z_j > \sum_{l=1}^{L} w_l \times \left(\frac{\alpha^n_{\langle i\rangle_n,l} - \alpha^n_{\langle 1\rangle_n,l}}{\sum_{k=1}^{K}\alpha^n_{k,l}}\right), i = 2,\cdots,K\right\}$$

$$= \frac{1}{\left(\sqrt{2\pi}\right)^{K-1}} \times \int\cdots\int_{\sum_{j=1}^{i-1}u_{j,i-1}z_j > \sum_{l=1}^{L} w_l\times\left(\frac{\alpha^n_{\langle i\rangle_n,l} - \alpha^n_{\langle 1\rangle_n,l}}{\sum_{k=1}^{K}\alpha^n_{k,l}}\right)} \prod_{k=1}^{K-1}\exp\left(-\frac{z_k^2}{2}\right)dz_{K-1}\cdots dz_1,$$

where Z_k, for $k = 1, 2, \cdots, K-1$, are independent standard normal random variables. The value of $V^N(\boldsymbol{S}^n)$ is the integration of the density of a $K-1$ dimensional standard normal distribution over an area covered by

$$\sum_{j=1}^{i-1} u_{j,i-1}z_j = \sum_{l=1}^{L} w_l \times \left(\frac{\alpha^n_{\langle i\rangle_n,l} - \alpha^n_{\langle 1\rangle_n,l}}{\sum_{k=1}^{K}\alpha^n_{k,l}}\right), \text{ for } i = 2,\cdots,K.$$

Following the approximation method in Peng et. al. (2018), we use the volume of the hypersphere centered at the origin to approximate the value of the integration over the entire region covered by the hyperplanes, i.e.,

$$V^N(\boldsymbol{S}^n) \approx \frac{\pi \times \min\left[\left(d_2^n\right)^2, \cdots, \left(d_K^n\right)^2\right]}{2} \triangleq \widetilde{V}^N(\boldsymbol{S}^n),$$

where d_k^n is the approximated distance from the origin to hyperplane, and

$$\left(d_k^n\right)^2 \triangleq \frac{\left[\sum_{l=1}^{L} w_l \times \left(\frac{\alpha^n_{\langle 1\rangle_n,l} - \alpha^n_{\langle k\rangle_n,l}}{\sum_{i=1}^{K}\alpha^n_{i,l}}\right)\right]^2}{\sum_{l=1}^{L}\left[w_l^2 \times \frac{\alpha^n_{\langle 1\rangle_n,l}\times\left(\sum_{i=1}^{K}\alpha^n_{i,l} - \alpha^n_{\langle 1\rangle_n,l}\right) + \alpha^n_{\langle k\rangle_n,l}\times\left(\sum_{i=1}^{K}\alpha^n_{i,l} - \alpha^n_{\langle k\rangle_n,l}\right) + 2\times\alpha^n_{\langle 1\rangle_n,l}\times\alpha^n_{\langle k\rangle_n,l}}{\left(\sum_{i=1}^{K}\alpha^n_{i,l}\right)^2\times\left(\sum_{i=1}^{K}\alpha^n_{i,l}+1\right)}\right]}$$

$$\approx \frac{\left[\sum_{l=1}^{L} w_l \times \left(\frac{\alpha^n_{\langle 1\rangle_n,l} - \alpha^n_{\langle k\rangle_n,l}}{\sum_{i=1}^{K}\alpha^n_{i,l}}\right)\right]^2}{\sum_{l=1}^{L}\left[w_l^2 \times \frac{\alpha^n_{\langle 1\rangle_n,l}\times\left(\sum_{i=1}^{K}\alpha^n_{i,l} - \alpha^n_{\langle 1\rangle_n,l}\right) + \alpha^n_{\langle k\rangle_n,l}\times\left(\sum_{i=1}^{K}\alpha^n_{i,l} - \alpha^n_{\langle k\rangle_n,l}\right) + 2\times\alpha^n_{\langle 1\rangle_n,l}\times\alpha^n_{\langle k\rangle_n,l}}{\left(\sum_{i=1}^{K}\alpha^n_{i,l}\right)^2\times\left(\sum_{i=1}^{K}\alpha^n_{i,l}+1\right)}\right] + \lambda}.$$

According to Eq. (18), we have

$$v_l^{KG}(\boldsymbol{S}^n) = E^n\left[V^N\left(T\left(\boldsymbol{S}^n, l, \boldsymbol{M}_l^{n+1}\right)\right)\right].$$

To calculate the value of $E^n\left[V^N\left(\boldsymbol{S}^{n+1}\right)|l^n = l\right] = E^n\left[V^N\left(T\left(\boldsymbol{S}^n, l, \boldsymbol{M}_l^{n+1}\right)\right)\right]$, we use the following approximation method described by Bertsekas (2005):

$$E^n\left[V^N\left(T\left(\boldsymbol{S}^n, l, \boldsymbol{M}_l^{n+1}\right)\right)\right] = E^n\left[V^N\left(\boldsymbol{S}^n + \boldsymbol{M}_l^{n+1}\boldsymbol{e}_l^L\right)\right] \approx V^N\left(\boldsymbol{S}^n + E^n\left[\boldsymbol{M}_l^{n+1}\right]\boldsymbol{e}_l^L\right).$$

With Assumption 1, the evaluation result with respect to criterion l satisfies the categorical distribution, i.e., $\boldsymbol{M}_l^{n+1}|\boldsymbol{p}_l \sim Cat(\boldsymbol{p}_l)$, allowing us to obtain

$$E^n\left[\boldsymbol{M}_l^{n+1}\right] = E\left[E\left[\boldsymbol{M}_l^{n+1}|\boldsymbol{P}_l\right]|\boldsymbol{S}^n\right] = E\left[\left(p_{1,l}, \cdots, p_{K,l}\right)^T|\boldsymbol{S}^n\right] = \left(\frac{\alpha_{1,l}^n}{\sum\limits_{i=1}^{K}\alpha_{i,l}^n}, \cdots, \frac{\alpha_{K,l}^n}{\sum\limits_{i=1}^{K}\alpha_{i,l}^n}\right)^T.$$

Define $\widetilde{\boldsymbol{S}}_l^{n+1} \triangleq \boldsymbol{S}^n + E^n\left[\boldsymbol{M}_l^{n+1}\right]\boldsymbol{e}_l^L$, where the elements of $\widetilde{\boldsymbol{S}}_l^{n+1} = \left(\widetilde{\boldsymbol{\alpha}}_1^{n+1}, \cdots, \widetilde{\boldsymbol{\alpha}}_L^{n+1}\right)$ can be calculated by

$$\widetilde{\boldsymbol{\alpha}}_l^{n+1} \triangleq \left(\widetilde{\alpha}_{1,l}^{n+1}, \cdots, \widetilde{\alpha}_{K,l}^{n+1}\right)^T = \boldsymbol{\alpha}_l^n + \left(\frac{\alpha_{1,l}^n}{\sum_{i=1}^{K}\alpha_{i,l}^n}, \cdots, \frac{\alpha_{K,l}^n}{\sum_{i=1}^{K}\alpha_{i,l}^n}\right)^T = \left(\alpha_{1,l}^n \times \frac{\sum_{i=1}^{K}\alpha_{i,l}^n + 1}{\sum_{i=1}^{K}\alpha_{i,l}^n}, \cdots, \alpha_{K,l}^n \times \frac{\sum_{i=1}^{K}\alpha_{i,l}^n + 1}{\sum_{i=1}^{K}\alpha_{i,l}^n}\right)^T$$

and $\widetilde{\boldsymbol{\alpha}}_i^{n+1} \triangleq \left(\widetilde{\alpha}_{1,i}^{n+1}, \cdots, \widetilde{\alpha}_{K,i}^{n+1}\right)^T = \boldsymbol{\alpha}_i^n$ for $i \neq l$.

In the state of $\widetilde{\boldsymbol{S}}_l^{n+1}$, the estimation of the global priority g_k is

$$E\left[g_k|\widetilde{\boldsymbol{S}}_l^{n+1}\right] = E\left[\sum_{d=1}^{L}\left(w_d \times \frac{\widetilde{\alpha}_{k,d}^{n+1}}{\sum\limits_{i=1}^{K}\widetilde{\alpha}_{i,d}^{n+1}}\right)\right] = \sum_{d=1,d\neq l}^{L}\left(w_d \times \frac{\alpha_{k,d}^n}{\sum\limits_{i=1}^{K}\alpha_{i,d}^n}\right) + w_l \times \frac{\widetilde{\alpha}_{k,l}^{n+1}}{\sum\limits_{i=1}^{K}\alpha_{i,l}^n + 1}$$

$$= \sum_{d=1}^{L}\left(w_d \times \frac{\alpha_{k,d}^n}{\sum\limits_{i=1}^{K}\alpha_{i,d}^n}\right) = g_k^n.$$

Thus, the estimated order statistics with respect to $\widetilde{\boldsymbol{S}}_l^{n+1}$ are equal to $\langle k\rangle_n$ for $k = 1, 2, \cdots, K$. Therefore,

$$E^n\left[V^N\left(T\left(\boldsymbol{S}^n, l, \boldsymbol{M}_l^{n+1}\right)\right)\right] \approx \widetilde{V}^N\left(\widetilde{\boldsymbol{S}}_l^{n+1}\right) = \frac{\pi \times \min\left[\left(d_{2,l}^{n+1}\right)^2, \cdots, \left(d_{K,l}^{n+1}\right)^2\right]}{2},$$

where

$$\left(d_{k,l}^{n+1}\right)^2 \triangleq \frac{\left[\sum_{d=1}^{L} w_d \times \left(\frac{\alpha_{\langle 1\rangle_n,d}^n - \alpha_{\langle k\rangle_n,d}^n}{\sum_{i=1}^{K}\alpha_{i,d}^n}\right)\right]^2}{\sum_{d=1}^{L}\left[w_d^2 \times \frac{\widetilde{\alpha}_{\langle 1\rangle_n,d}^{n+1} \times \left(\sum_{i=1}^{K}\widetilde{\alpha}_{i,d}^{n+1} - \widetilde{\alpha}_{\langle 1\rangle_n,d}^{n+1}\right) + \widetilde{\alpha}_{\langle k\rangle_n,d}^{n+1} \times \left(\sum_{i=1}^{K}\widetilde{\alpha}_{i,d}^{n+1} - \widetilde{\alpha}_{\langle k\rangle_n,d}^{n+1}\right) + 2 \times \widetilde{\alpha}_{\langle 1\rangle_n,d}^{n+1} \times \widetilde{\alpha}_{\langle k\rangle_n,d}^{n+1}}{\left(\sum_{i=1}^{K}\widetilde{\alpha}_{i,d}^{n+1}\right)^2 \times \left(\sum_{i=1}^{K}\widetilde{\alpha}_{i,d}^{n+1} + 1\right)}\right]},$$

and $\widetilde{\boldsymbol{\alpha}}_l^{n+1} \triangleq \left(\widetilde{\alpha}_{1,l}^{n+1}, \cdots, \widetilde{\alpha}_{K,l}^{n+1}\right)^T = \left(\alpha_{1,l}^n \times \frac{\sum_{i=1}^{K}\alpha_{i,l}^n + 1}{\sum_{i=1}^{K}\alpha_{i,l}^n}, \cdots, \alpha_{K,l}^n \times \frac{\sum_{i=1}^{K}\alpha_{i,l}^n + 1}{\sum_{i=1}^{K}\alpha_{i,l}^n}\right)^T$, $\widetilde{\boldsymbol{\alpha}}_i^{n+1} \triangleq \left(\widetilde{\alpha}_{1,i}^{n+1}, \cdots, \widetilde{\alpha}_{K,i}^{n+1}\right)^T = \boldsymbol{\alpha}_i^n$ for $i \neq l$.

In summary, the AHP-KG factor for criterion l can be approximated by

$$v_l^{KG}(\boldsymbol{S}^n) = E^n\left[V^N\left(T\left(\boldsymbol{S}^n, l, \boldsymbol{M}_l^{n+1}\right)\right)\right] \approx \widetilde{V}^N\left(\widetilde{\boldsymbol{S}}_l^{n+1}\right) = \frac{\pi \times \min\left[\left(d_{2,l}^{n+1}\right)^2, \cdots, \left(d_{K,l}^{n+1}\right)^2\right]}{2} \triangleq \widetilde{v}_l^{KG}(\boldsymbol{S}^n).$$

And the AHP-KG allocation rule is

$$l^{KG}(\boldsymbol{S}^n) \in \underset{l\in\{1,\cdots,L\}}{\operatorname{argmax}}\,\widetilde{v}_l^{KG}(\boldsymbol{S}^n) = \underset{l\in\{1,\cdots,L\}}{\operatorname{argmax}}\left\{\min\left[\left(d_{2,l}^{n+1}\right)^2, \cdots, \left(d_{K,l}^{n+1}\right)^2\right]\right\}.$$

■

Appendix B: Proof of Theorem 2

Proof. For any finite terminal time N, we define η_l^N for $l = 1, 2, \cdots, L$ as the number of times that criterion l is evaluated. Furthermore, η_l^∞ for $l = 1, 2, \cdots, L$ denote the limit of η_l^N as N approaches infinity; namely,

$$\eta_l^N \triangleq \sum_{n=0}^{N-1}\mathbf{1}\{l^n = l\} \text{ and } \eta_l^\infty \triangleq \lim_{N\to\infty}\eta_l^N.$$

As mentioned above, the terminal value of the state $\boldsymbol{S}^n$ can be approximated by

$$\widetilde{V}^N(\boldsymbol{S}^n) = \frac{\pi \times \min\left[\left(d_2^n\right)^2, \cdots, \left(d_K^n\right)^2\right]}{2},$$

where

$$\left(d_k^n\right)^2 \overset{\Delta}{=} \frac{\left[\sum_{l=1}^{L} w_l \times \left(\frac{\alpha^n_{\langle 1\rangle_n,l}-\alpha^n_{\langle k\rangle_n,l}}{\sum_{i=1}^{K}\alpha^n_{i,l}}\right)\right]^2}{\sum_{l=1}^{L}\left[w_l^2 \times \frac{\alpha^n_{\langle 1\rangle_n,l}\times\left(\sum_{i=1}^{K}\alpha^n_{i,l}-\alpha^n_{\langle 1\rangle_n,l}\right)+\alpha^n_{\langle k\rangle_n,l}\times\left(\sum_{i=1}^{K}\alpha^n_{i,l}-\alpha^n_{\langle k\rangle_n,l}\right)+2\times\alpha^n_{\langle 1\rangle_n,l}\times\alpha^n_{\langle k\rangle_n,l}}{\left(\sum_{i=1}^{K}\alpha^n_{i,l}\right)^2\times\left(\sum_{i=1}^{K}\alpha^n_{i,l}+1\right)}\right]}.$$

At each time $0 \le n < N$, we choose the criterion with the largest approximate AHP-KG factor $\widetilde{v}_l^{KG}(\boldsymbol{S}^n) = \frac{\pi\times\left\{min\left[\left(d_{2,l}^{n+1}\right)^2,\cdots,\left(d_{K,l}^{n+1}\right)^2\right]\right\}}{2}$ to evaluate next. The difference among $\left(d_{k,l}^{n+1}\right)^2$ and $\left(d_k^n\right)^2$ is caused by the

$$\frac{\widetilde{\alpha}^{n+1}_{\langle 1\rangle_n,l}\times\left(\sum_{i=1}^{K}\widetilde{\alpha}^{n+1}_{i,l}-\widetilde{\alpha}^{n+1}_{\langle 1\rangle_n,l}\right)+\widetilde{\alpha}^{n+1}_{\langle k\rangle_n,l}\times\left(\sum_{i=1}^{K}\widetilde{\alpha}^{n+1}_{i,l}-\widetilde{\alpha}^{n+1}_{\langle k\rangle_n,l}\right)+2\times\widetilde{\alpha}^{n+1}_{\langle 1\rangle_n,l}\times\widetilde{\alpha}^{n+1}_{\langle k\rangle_n,l}}{\left(\sum_{i=1}^{K}\widetilde{\alpha}^{n+1}_{i,l}\right)^2\times\left(\sum_{i=1}^{K}\widetilde{\alpha}^{n+1}_{i,l}+1\right)}$$

$$=\frac{\frac{\alpha^n_{\langle 1\rangle_n,l}}{\sum_{i=1}^{K}\alpha^n_{i,l}}\times\left(1-\frac{\alpha^n_{\langle 1\rangle_n,l}}{\sum_{i=1}^{K}\alpha^n_{i,l}}\right)+\frac{\alpha^n_{\langle k\rangle_n,l}}{\sum_{i=1}^{K}\alpha^n_{i,l}}\times\left(1-\frac{\alpha^n_{\langle k\rangle_n,l}}{\sum_{i=1}^{K}\alpha^n_{i,l}}\right)+2\times\frac{\alpha^n_{\langle 1\rangle_n,l}}{\sum_{i=1}^{K}\alpha^n_{i,l}}\times\frac{\alpha^n_{\langle k\rangle_n,l}}{\sum_{i=1}^{K}\alpha^n_{i,l}}}{\left(\sum_{i=1}^{K}\alpha^n_{i,l}+2\right)}$$

$$=\frac{\alpha^n_{\langle 1\rangle_n,l}\times\left(\sum_{i=1}^{K}\alpha^n_{i,l}-\alpha^n_{\langle 1\rangle_n,l}\right)+\alpha^n_{\langle k\rangle_n,l}\times\left(\sum_{i=1}^{K}\alpha^n_{i,l}-\alpha^n_{\langle k\rangle_n,l}\right)+2\times\alpha^n_{\langle 1\rangle_n,l}\times\alpha^n_{\langle k\rangle_n,l}}{\left(\sum_{i=1}^{K}\alpha^n_{i,l}\right)^2\times\left(\sum_{i=1}^{K}\alpha^n_{i,l}+2\right)}$$

$$<\frac{\alpha^n_{\langle 1\rangle_n,l}\times\left(\sum_{i=1}^{K}\alpha^n_{i,l}-\alpha^n_{\langle 1\rangle_n,l}\right)+\alpha^n_{\langle k\rangle_n,l}\times\left(\sum_{i=1}^{K}\alpha^n_{i,l}-\alpha^n_{\langle k\rangle_n,l}\right)+2\times\alpha^n_{\langle 1\rangle_n,l}\times\alpha^n_{\langle k\rangle_n,l}}{\left(\sum_{i=1}^{K}\alpha^n_{i,l}\right)^2\times\left(\sum_{i=1}^{K}\alpha^n_{i,l}+1\right)}.$$

Therefore,

$$\left(d_{k,l}^{n+1}\right)^2=\frac{\left[\sum_{d=1}^{L} w_d \times \left(\frac{\alpha^n_{\langle 1\rangle_n,d}-\alpha^n_{\langle k\rangle_n,d}}{\sum_{i=1}^{K}\alpha^n_{i,d}}\right)\right]^2}{\sum_{d=1}^{L}\left[w_d^2 \times \frac{\widetilde{\alpha}^{n+1}_{\langle 1\rangle_n,d}\times\left(\sum_{i=1}^{K}\widetilde{\alpha}^{n+1}_{i,d}-\widetilde{\alpha}^{n+1}_{\langle 1\rangle_n,d}\right)+\widetilde{\alpha}^{n+1}_{\langle k\rangle_n,d}\times\left(\sum_{i=1}^{K}\widetilde{\alpha}^{n+1}_{i,d}-\widetilde{\alpha}^{n+1}_{\langle k\rangle_n,d}\right)+2\times\widetilde{\alpha}^{n+1}_{\langle 1\rangle_n,d}\times\widetilde{\alpha}^{n+1}_{\langle k\rangle_n,d}}{\left(\sum_{i=1}^{K}\widetilde{\alpha}^{n+1}_{i,d}\right)^2\times\left(\sum_{i=1}^{K}\widetilde{\alpha}^{n+1}_{i,d}+1\right)}\right]}$$

$$>\frac{\left[\sum_{d=1}^{L} w_d \times \left(\frac{\alpha^n_{\langle 1\rangle_n,d}-\alpha^n_{\langle k\rangle_n,d}}{\sum_{i=1}^{K}\alpha^n_{i,d}}\right)\right]^2}{\sum_{d=1}^{L}\left[w_d^2 \times \frac{\alpha^n_{\langle 1\rangle_n,d}\times\left(\sum_{i=1}^{K}\alpha^n_{i,d}-\alpha^n_{\langle 1\rangle_n,d}\right)+\alpha^n_{\langle k\rangle_n,d}\times\left(\sum_{i=1}^{K}\alpha^n_{i,d}-\alpha^n_{\langle k\rangle_n,d}\right)+2\times\alpha^n_{\langle 1\rangle_n,d}\times\alpha^n_{\langle k\rangle_n,d}}{\left(\sum_{i=1}^{K}\alpha^n_{i,d}\right)^2\times\left(\sum_{i=1}^{K}\alpha^n_{i,d}+1\right)}\right]}$$

$$=\left(d_k^n\right)^2.$$

And, we have $\widetilde{v}_l^{KG}(\boldsymbol{S}^n) > \widetilde{V}^N(\boldsymbol{S}^n)$ for every $l \in \{1,2,\cdots,L\}$.

As given by Eq. (5), $\sum_{k=1}^{K}\alpha^n_{i,l} = \sum_{k=1}^{K}\alpha^0_{i,l} + \eta^n_l$. When $\eta^n_l \to \infty$, we have $\left(d_{k,l}^{n+1}\right)^2 \to \left(d_k^n\right)^2$. Therefore, $\widetilde{v}_l^{KG,n} - \widetilde{V}^N(\boldsymbol{S}^n) \to 0$ as $\eta^n_l \to \infty$.

As our total number of evaluations $N \to \infty$, for at least one criterion $l \in \{1,2,\cdots,L\}$, we must have $\eta^N_l \to \infty$ and therefore $lim_{N\to\infty}\widetilde{v}_l^{KG}(\boldsymbol{S}^N) - \widetilde{V}^N(\boldsymbol{S}^N) = 0$. Let F denote the set of criteria for which η^N_f remains finite. Eventually, we must reach a state where we allocate no further evaluation to F. According to the derivation above, we have $\widetilde{v}_f^{KG}(\boldsymbol{S}^N) > \widetilde{V}^N(\boldsymbol{S}^N)$ for $f \in F$. Then, there exists some number of evaluations N' such that $\widetilde{v}_l^{KG}(\boldsymbol{S}^n) < \widetilde{v}_f^{KG}(\boldsymbol{S}^n)$ for $n \ge N'$ and $l \in \{1,2,\cdots,L\} - F$. In that case, the AHP-KG policy will allocate the next evaluation to F. Thus, by contradiction, F is empty.

With an infinite evaluation budget, each criterion will have an infinite number of experts allocated, that is, $\eta^\infty_l \to \infty$ for $l = 1,2,\cdots,L$. The Central Limit Theorem justifies the AHP-KG policy's asymptotic optimality.■

References

Aguarón, J., Escobar, M. T., & Moreno-Jiménez, J. M. (2021). Reducing inconsistency measured by the geometric consistency index in the analytic hierarchy process. *European Journal of Operational Research, 288*(2), 576–583.

Altuzarra, A., Moreno-Jiménez, J. M., & Salvador, M. (2007). A Bayesian priorization procedure for AHP-group decision making. *European Journal of Operational Research, 182*(1), 367–382.

Bard, J. F., & Sousk, S. F. (1990). A trade-off analysis for rough terrain cargo handlers using the AHP: An example of group decision making. *IEEE Transaction on Engineering Management, 37*(3), 222–228.

Bertsekas, D. P. (2005). Dynamic programming and suboptimal control: A survey from ADP to MPC. *European Journal of Control, 11*(4), 310–334.

Brunelli, M. (2014). *Introduction to the analytic hierarchy process*. Springer.

Eckman, D. J., & Henderson, S. G. (2022). Posterior-Based stopping rules for bayesian ranking-and-selection procedures. *INFORMS Journal on Computing, 34*(3), 1305–1840.

Fan, W., Hong, L. J., & Zhang, X. (2020). Distributionally robust selection of the best. *Management Science, 66*(1), 190–208.

Frazier, P. I., Powell, W. B., & Dayanik, S. (2008). A knowledge-gradient policy for sequential information collection. *SIAM Journal on Control and Optimization, 47*(5), 2410–2439.

Gao, S., Xiao, H., Chen, W., & Zhou, E. (2017). Robust ranking and selection with optimal computing budget allocation. *Automatica, 81*, 30–36.

Gargallo, P., Moreno-Jiménez, J. M., & Salvador, M. (2007). AHP-Group decision making: A Bayesian approach based on mixtures for group pattern identification. *Group Decision and Negotiation, 16*, 485–506.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis*. New York: Taylor & Francis Group.

Groves, M., & Branke, J. (2019). Top-k selection with pairwise comparisons. *European Journal of Operational Research, 274*(2), 615–626.

Ho, W. (2008). Integrated analytic hierarchy process and its applications–A literature review. *European Journal of Operational Research, 186*(1), 211–228.

Ho, W., & Ma, X. (2018). The state-of-the-art integrations and applications of the analytic hierarchy process. *European Journal of Operational Research, 267*(2), 399–414.

Huang, E., Zhang, S., Lee, L. H., Chew, E. P., & Chen, C. (2015). Improving analytic hierarchy process expert allocation using optimal computing budget allocation. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 46*(8), 1140–1147.

Ishizaka, A., & Labib, A. (2011). Review of the main developments in the analytic hierarchy process. *Expert Systems with Applications, 38*(11), 14336–14345.

Jalao, E. R., Wu, T., & Shunk, D. (2014). An intelligent decomposition of pairwise comparison matrices for large-scale decisions. *European Journal of Operational Research, 238*(1), 270–280.

Kaminski, B., & Szufel, P. (2014). Asynchronous knowledge gradient policy for ranking and selection. In *Proceedings of the 2014 winter simulation conference* (pp. 3785–3796).

Kokangül, A., Polat, U., & Dağsuyu, C. (2017). A new approximation for risk assessment using the AHP and Fine Kinney methodologies. *Safety Science, 91*, 24–32.

Kou, G., Ergu, D., Lin, C., & Chen, Y. (2016). Pairwise comparison matrix in multiple criteria decision making. *Technological and Economic Development of Economy, 22*(5), 738–765.

Kou, G., Xiao, H., Cao, M., & Lee, L. H. (2021). Optimal computing budget allocation for the vector evaluated genetic algorithm in multi-objective simulation optimization. *Automatica*, 129.

Lee, L. H., Chew, E. P., Teng, S., & Goldsman, D. (2010). Finding the non-dominated Pareto set for multi-objective simulation models. *IIE Transactions, 42*(9), 656–674.

Lin, C., & Kou, G. (2015). Bayesian revision of the individual pair-wise comparison matrices under consensus in AHP-GDM. *Applied Soft Computing, 35*, 802–811.

Li, H., Xu, X., Peng, Y., & Chen, C.H. (2019). Efficient sampling for selecting important nodes in random network. arXiv preprint arXiv:1901.03466.

Li, J., Liu, W., Pedrielli, G., Lee, L. H., & Chew, E. P. (2018). Optimal computing budget allocation to select the non-dominated systems - a large deviations perspective. *IEEE Transactions on Automatic Control, 63*(9), 2913–2927.

Lyu, H. M., Zhou, W. H., Shen, S. L., & Zhou, A. N. (2020). Inundation risk assessment of metro system using AHP and TFN-AHP in Shenzhen. *Sustainable Cities and Society, 56*, Article 102103.

Ng, K. W., Tian, G., & Tang, M. (2011). *Dirichlet and related distribution: Theory, methods and application*. New York: John Wiley & Sons.

Peng, Y., Chong, E. K., Chen, C. H., & Fu, M. C. (2018). Ranking and selection as stochastic control. *IEEE Transactions on Automatic Control, 63*(8), 2359–2373.

Ping, P., Wang, K., Kong, D., & Chen, G. (2018). Estimating probability of success of escape, evacuation, and rescue (EER) on the offshore platform by integrating Bayesian network and fuzzy AHP. *Journal of Loss Prevention in the Process Industries, 54*, 57–68.

Powell, W. B., & Ryzhov, I. O. (2012). *Optimal learning*. USA: John Wiley & Sons.

Saaty, T. L. (1980). *The analytic hierarchy process*. New York: Mc-Graw-Hill.

Saaty, T. L. (1994). *Fundamentals of decision making and priority theory with the analytic hierarchy process*. RWS publications.

Saaty, T. L., & Vargas, L. G. (2012). *Models, methods, concepts & applications of the analytic hierarchy process*. USA: Springer.

Subramanian, N., & Ramanathan, R. (2012). A review of applications of analytic hierarchy process in operations management. *International Journal of Production Economics, 138*(2), 215–241.

Wang, H., Kou, G., & Peng, Y. (2021). An iterative algorithm to derive priority from large-scale sparse pairwise comparison matrix. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 52*(5), 3038–3051.

Wu, J., & Frazier, P. I. (2016). The parallel knowledge gradient method for batch Bayesian optimization. In *30th Conference on neural information processing systems*.

Xiao, H., Chen, H., & Lee, L. H. (2019). An efficient simulation procedure for ranking the top simulated designs in the presence of stochastic constraints. *Automatica, 103*, 106–115.

Xiao, H., & Gao, S. (2018). Simulation budget allocation for selecting the top-m designs with input uncertainty. *IEEE Transactions on Automatic Control, 63*(9), 3127–3134.

Zhang, S., Xu, J., Lee, L. H., Chew, E. P., Wong, W. P., & Chen, C. H. (2016). Optimal computing budget allocation for particle swarm optimization in stochastic optimization. *IEEE Transactions on Evolutionary Computation, 21*(2), 206–219.

Zhong, Y., & Hong, L. J. (2021). Knockout-tournament procedures for large-scale ranking and selection in parallel computing environments. *Operations Research, 70*(1), 432–453.