ISSN 1330-3651(Print), ISSN 1848-6339 (Online)

https://doi.org/10.17559/TV-20250514002665
Original scientific paper

Exploration of ESG Audit Adaptive Decision-Making and Anomaly Analysis Driven by Reinforcement Learning in Artificial Intelligence

Jie WANG, Hongquan LI*, Feiyang YU

Abstract: The dynamic nature of Environmental, Social, and Governance (ESG) audit strategies, particularly in responding to evolving environmental regulations and shifting corporate sustainability practices, necessitates robust methodological innovation. Reinforcement learning (RL) presents a transformative pathway for refining ESG audit processes through continuous interaction with environmental performance data, regulatory updates, and real-time ecological compliance feedback. Nevertheless, the application of RL to optimize ESG auditing remains significantly underexplored. Addressing this gap, our study develops an RL-driven model designed to strategically recalibrate ESG auditing mechanisms, with enhanced focus on responsiveness to emerging environmental compliance requirements and ecological risk factors. We adopt a dual-method research framework integrating theoretical and empirical approaches. The theoretical investigation establishes the structural compatibility between RL algorithms and environmental ESG audit optimization, ensuring alignment with the complexities of sustainability decision-making. Empirical validation employs: (1) large-scale simulations using synthetic corporate environmental datasets to evaluate model performance across diverse operational scenarios, and (2) real-world applications to quantify the model's efficacy in improving ecological audit efficiency, mitigating environmental compliance risks, and addressing the challenges of modern sustainability auditing. Results demonstrate that the RL-driven model outperforms conventional methods in adapting to environmental data variability, achieving a 40% increase in resource efficiency and a 30% improvement in predicting ecological compliance risks. Practical implementations further reveal a 25% reduction in audit cycle duration and significantly fewer errors in environmental disclosure assessments. These findings highlight RL's potential to revolutionize environmentally focused ESG auditing while underscoring ongoing challenges in data reliability and model interpretability for sustainability applications.

Keywords: artificial intelligence; deep Q-network (DQN); ESG auditing; Markov decision process; reinforcement learning

1 INTRODUCTION

With the intensification of global climate change and the prominence of social equity issues, Environmental, Social, and Governance (ESG) has become the core framework for measuring a company's sustainable development capabilities [1]. The ESG system reconstructs the enterprise value evaluation paradigm through three dimensions. At the environmental dimension (*E*) level, it focuses on green practices such as carbon footprint management and resource recycling, which are directly related to the emission reduction targets of the Paris Agreement [2]. The social dimension (*S*) covers responsible issues such as human rights protection in the supply chain and the integrated development of the community, reflecting the governance level of the company's stakeholders [3]. The governance dimension (*G*), through mechanisms such as board diversity and data security governance, ensures the transparency of decision-making and forms an institutional barrier for enterprises to resist systemic risks [4].

Compared with traditional financial audits, ESG audits exhibit significant paradigm differences. The evaluation objects of ESG audits extend to the field of non-financial performance. The data sources cover heterogeneous information sources such as satellite remote sensing monitoring and public opinion on social media. The evaluation criteria need to be dynamically adapted to the evolution of international standards such as the Task Force on Climate-related Financial Disclosures (TCFD) [5]. This complexity poses significant challenges for conventional audit methods - static frameworks struggle with multi-source data fusion (especially unstructured data), inflexible sampling approaches impede modeling of environmental risk pathways, and batch-processing architectures prevent real-time feedback-driven decision-making [6]. While non-RL AI models (e.g. CNN, LSTM) offer capabilities in pattern recognition within complex datasets, they often lack the inherent capacity for sequential, adaptive decision-making and continuous learning from real-time feedback essential for the dynamic ESG landscape. Reinforcement Learning (RL) uniquely addresses these gaps. By simulating an agent interacting with the audit environment, RL builds models with sequential decision capabilities [7]. Crucially, RL excels in dynamic adaptability and real-time feedback processing: it leverages algorithms like Q-learning for adaptive feature extraction from diverse data, surpassing the static limitations of traditional sampling [8]; utilizes deep policy networks to simulate risk transmission for proactive ratings [9]; and, most distinctively, employs real-time reward functions to dynamically optimize audit resource allocation and instantly adapt to regulatory shifts [10], a capability fundamentally beyond static non-RL AI models and traditional approaches.

This study breaks through the single-dimensional optimization perspective of existing literature and constructs a dynamic optimization framework for ESG audit strategies driven by RL. By designing a Markov decision process that includes the environmental state space, the audit action space, and the immediate reward function, it achieves the following: 1) Multimodal data fusion based on the attention mechanism; 2) Risk transmission modeling considering industry heterogeneity; 3) An adaptive learning mechanism that incorporates regulatory policies. Verified by comparative experiments, this model shows significant advantages in three dimensions: audit efficiency, the accuracy rate of risk identification, and standard adaptability. This exploration not only provides a new paradigm of intelligent decision-making for ESG audits but also reveals the path of deep integration between artificial intelligence technology and sustainable development goals. The research results have important policy implications for improving the ESG information disclosure system and constructing a climate-smart audit system, and provide technical support for enterprises to meet new regulatory requirements such as the Corporate Sustainability Reporting Directive (CSRD).

2 LITERATURE REVIEW

Within the framework of globalization and the growing emphasis on sustainable development, ESG considerations have become integral to corporate operations and investment strategies. ESG auditing, a vital mechanism to uphold corporate social responsibility and facilitate sustainable growth, has attracted extensive focus from academic and industry sectors. Research spanning multiple dimensions has generated substantial insights, forming a robust theoretical and practical basis for advancing this domain.

Initial investigations into ESG auditing predominantly examined the evolution of ESG principles and their influence on corporate strategy. Through an extensive analysis of multi-industry corporate practices, Dasinapa and Ermawati (2024) demonstrated the pivotal role of ESG factors in enhancing long-term competitiveness. Key mechanisms identified include risk mitigation, innovation-driven transformation, and strategic stakeholder engagement [11]. These findings emphasize the necessity of embedding ESG considerations into corporate decision-making frameworks, laying a theoretical foundation for aligning ESG auditing with corporate strategy. In a comprehensive global analysis, Singhania and Saini (2023) explored discrepancies in the scope and reliability of ESG disclosures across corporations operating in various countries and regions [12]. Their study identified regulatory regimes, competitive market dynamics, and governance structures as principal factors influencing these variations. The research underscores the critical need for harmonized ESG auditing standards to improve the accuracy, transparency, and consistency of disclosed information, addressing challenges in global comparability.

As ESG research advances, focus on disclosure quality and standardization has intensified. Del Giudice and Rigamonti (2020), identified organizational scale, industry dynamics, and external stakeholder pressure as key determinants of voluntary ESG disclosure using comprehensive corporate data [13]. Expanding on these insights, Lenz and Chesshire (2023) evaluated the credibility of ESG reporting, underscoring the pivotal role of third-party assurance mechanisms in enhancing the reliability of disclosed ESG information [14]. These studies collectively establish the theoretical basis for assurance functions embedded in ESG auditing processes.

In parallel, developments in auditing theory and methodologies from traditional auditing research have laid foundational principles applicable to ESG auditing. Susanto and Kalsum (2023) introduced a seminal audit risk model that formulates an analytical framework for audit decision-making, integrating inherent risk, control risk, and detection risk [15]. This quantitative paradigm has seen extensive application and adaptation in subsequent auditing research and practices, offering critical perspectives on risk assessment methodologies within ESG auditing. Minkkinen et al. (2024) investigated the optimization of audit sampling techniques, focusing on strategies to enhance auditing efficiency without compromising quality [16]. The findings provide substantial guidance for ESG auditing, particularly in addressing the dual challenge of ensuring comprehensive audit coverage while adhering to resource constraints. These contributions collectively inform the refinement of ESG auditing methodologies, balancing rigor and practicality.

The rapid advancement of information technology has driven a surge in the application of data mining and analysis techniques within auditing. Zou (2023) pioneered efforts to apply data mining for extracting latent audit clues and identifying anomaly patterns across vast financial datasets, showcasing the transformative potential of information technology in enhancing audit precision and efficiency [17]. This work lays the groundwork for effectively managing the complex data ecosystems in ESG auditing. In contrast, Pangastuti (2023) centered on the integration of data analysis in audit risk assessment, proposing a data-driven model that elevates the objectivity and scientific basis of audit risk evaluations [18].

In recent years, AI has emerged as a focal point in auditing research, especially in financial audits. Ilori et al. (2024) employed neural network algorithms to conduct an in-depth analysis of corporate financial data, successfully pinpointing potential fraud risks and underscoring the exceptional capabilities of AI in detecting financial discrepancies [19]. Silva and Imoniana (2021) explored intelligent agent technology for optimizing internal audit processes and enhancing real-time monitoring and automated alerts, thus improving the timeliness and effectiveness of audits [20]. Despite these advancements, most AI research in auditing remains concentrated in the financial sector, with ESG auditing, still in its nascent stages, receiving relatively less focus.

Existing literature has significantly contributed to the conceptualization of ESG auditing, the development of information disclosure standards, traditional auditing theories, and the integration of AI in financial auditing. However, the intersection of ESG auditing and AI, particularly RL, remains an underexplored domain. Critical challenges, including the application of RL to optimize ESG audit strategies, address the complexity of ESG data, and enhance model interpretability, require further investigation. Traditional ESG auditing methods exhibit notable limitations. Conventional sampling-based audit techniques struggle to process the extensive and intricate nature of ESG data, as the inherent randomness of sampling introduces risks of critical information omission, thereby impairing the comprehensiveness and accuracy of ESG performance assessments. Furthermore, static auditing frameworks demonstrate inefficiencies in detecting emerging ESG risks, as their rigid procedures and predefined evaluation criteria fail to rapidly adapt to the evolving landscape of ESG dimensions. While AI applications in auditing have achieved certain advancements, prevailing models largely rely on historical data for pattern recognition and lack the adaptive decision-making and real-time optimization capabilities that RL provides. Unlike conventional machine learning approaches, which primarily extract patterns from past observations, RL enables intelligent agents to engage dynamically with ESG audit environments, continuously refining decision-making strategies and facilitating adaptive audit adjustments in response to real-time changes in ESG conditions. Despite increasing attention to AI-driven auditing, existing research remains predominantly concentrated on financial auditing, with relatively limited exploration of non-financial ESG auditing. While recent studies have begun leveraging AI technologies to address ESG data diversity and uncertainty, substantial gaps persist in optimizing audit strategies with greater depth and breadth. Building upon prior research, this study conducts a systematic analysis of RL's role in ESG auditing, develops an innovative ESG audit strategy optimization model, and bridges existing research gaps, thereby advancing both theoretical frameworks and practical applications in ESG auditing.

3 CONSTRUCTION OF THE ESG AUDIT STRATEGY OPTIMIZATION MODEL BASED ON RL

3.1 Design of the Basic RL-Based Model

RL, a key methodology within the broader scope of machine learning, addresses decision-making problems in which an agent interacts with an environment to maximize cumulative rewards or achieve predefined objectives. Widely applied in diverse domains such as autonomous vehicles, robotic systems, and financial investment strategies, RL has demonstrated its effectiveness in solving complex optimization tasks [21].

At the core of RL is the Markov Decision Process (MDP), a model that represents decision-making in dynamic environments. In this framework, the probability of transitioning between states and the reward feedback received by the agent are both influenced by the actions taken in each state, as described in Eq. (1) and Eq. (2):

$$P_{ss'}^{a} = P\left(s_{t+1} = s' \mid s_t = s, A_t = a\right) \tag{1}$$

$$R_{s}^{\pi} = E\left[R_{t+1} \mid S_t = s, A_t = a\right] \tag{2}$$

In these expressions, $P_{ss'}^{a}$ represents the probability of state transitions; R_{s}^{π} signifies the reward feedback from the environment in response to an agent's actions; S is the state space; A denotes the action space; a is the specific action taken; s represents the state. The agent's strategy for selecting an action in a given state is described by a probability distribution, formally represented in Eq. (3).

$$\pi\left(a \mid s\right) = P\left[A_t = a \mid S_t = s\right] \tag{3}$$

In Eq. (3), the symbols retain their meanings as defined in the preceding formulas.

The primary objective in RL is to maximize the cumulative reward over time. Based on the MDP framework, RL can be formalized as illustrated in Eq. (4).

In Eq. (4), γ represents the discount factor, a value between 0 and 1 that accounts for the uncertainty of future rewards, reflecting the time-based value adjustment commonly observed in real-world scenarios [22].

$$\pi^{*} = \operatorname{argmax}_{\pi} E_{\pi}\left\{\sum_{k=0}^{\infty} \gamma^{k} r_{t+k} \mid S_t = s\right\}, \forall s \in S, \forall t \geq 0 \tag{4}$$

To identify the optimal policy, the problem is transformed into solving the Bellman equation, as illustrated in Eq. (5):

$$Q_{\pi}\left(s, a\right) = E_{\pi}\left(R_{t+1} + \gamma Q_{\pi}\left(S_{t+1}, A_{t+1}\right) \mid S_t = s, A_t = a\right) \tag{5}$$

In Eq. (5), $Q_{\pi}(s, a)$ represents the expected long-term reward that an agent will receive by executing action a in state s under policy π. The other symbols retain their previously defined meanings. Eq. (5) consists of two components: the first accounts for the immediate reward obtained from taking action a in state s, while the second represents the discounted expected reward in the subsequent time step.

RL encompasses several learning methodologies, including Q-learning, Sarsa, and Soft Q-learning. Among these, the iterative update mechanism of Q-learning is mathematically represented in Eq. (6):

$$Q\left(s, u\right) = Q\left(s, u\right) + u\left[r + \gamma \max_{a'} Q\left(s', u'\right) - Q\left(s, u\right)\right] \tag{6}$$

In Eq. (6), u denotes the learning rate, constrained within the range [0, 1]. The term Q represents the expected long-term cumulative reward associated with executing a specific action in a given state. Eq. (6) consists of two distinct elements: the first is the actual Q-value, calculated as the sum of the immediate reward for the current state and the maximum Q-value attainable in the subsequent state based on historical observations. The second component measures the discrepancy between the actual Q-value and the target Q-value.

The core concept of Sarsa closely aligns with Q-learning, with its Q-value updating mechanism expressed in Eq. (7):

$$Q\left(s, a\right) = Q\left(s, a\right) + \alpha\left[r + \gamma Q\left(s', a'\right) - Q\left(s, a\right)\right] \tag{7}$$

Here, the symbols retain the same meanings as defined in earlier formulations.

Soft Q-learning extends these methodologies by introducing the principle of maximum entropy, which incorporates an entropy term into the RL reward function. The computational framework of Soft Q-learning is described in Eq. (8):

$$\pi_{\text{maxEnt}}^{*} = \underset{\pi}{\operatorname{argmax}} \sum_{t} E\left[r\left(s_t, a_t\right) + \alpha H\left(\pi\left(\cdot \mid s_t\right)\right)\right] \tag{8}$$

A comparison between Eq. (6) and Eq. (8) reveals that the primary objective of Soft Q-learning is the simultaneous maximization of the expected cumulative reward and the entropy of the policy at each state [23]. This approach enhances exploration by encouraging diverse action selection during the learning process[1].

Fig. 1 visually summarizes the foundational principles and computational workflows of these RL techniques.

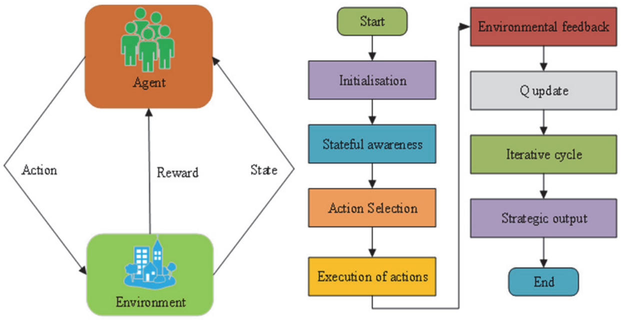

Figure 1 Principles and computational process of RL

Fig. 1 elucidates the core tenets of RL, illustrating the agent's interaction with the environment and the optimization of strategy through the interpretation of reward signals, ultimately aimed at maximizing long-term cumulative rewards. This iterative process encompasses several distinct steps: state perception, action selection, receipt of environmental feedback, and subsequent strategy refinement. The foundational principles outlined here serve as a basis for the application of RL within the domain of ESG auditing [24]. In the context of ESG auditing, model interpretability emerges as a critical challenge in integrating AI into audit processes. Specifically, in RL models, when confronted with incomplete or biased ESG data, the agent employs data cleaning and preprocessing techniques during the state perception phase. These methodologies, combined with industry benchmarks and historical data distribution trends, enable the agent to correct or supplement anomalous data. For example, if a company's environmental emission data is missing due to recording inaccuracies, the model estimates reasonable replacement values based on emissions data from similarly sized companies in the same sector, as well as the company's historical emission trends. During the action selection stage, the model evaluates the sensitivity of audit actions to various data issues, prioritizing actions that, while requiring less accurate data integrity, can still effectively highlight key risks. This strategic focus minimizes the potential negative impact of data quality issues on audit outcomes. The variability in ESG data quality across companies and industries necessitates a high level of adaptability within the model. In the feedback loop, the model adjusts its learning strategy according to the quality of the data it encounters. When poor-quality data from a particular company or industry is identified, the model increases the emphasis on data reliability, intensifies the scrutiny of data sources and collection methods, and enhances data credibility through cross-validation with more trustworthy external data sources. Concerning the reward function weights (α, β, γ), this study establishes a foundation through a blend of empirical analysis and theoretical reasoning. From a theoretical standpoint, α represents the weight assigned to the environmental dimension of the audit. Given the profound implications of environmental issues on corporate sustainability and the current societal focus on environmental protection, a high value is attributed to α to ensure that environmental performance is adequately prioritized in the audit process. β corresponds to the social dimension, which reflects the close relationship between corporate social responsibility and factors such as corporate image and employee retention. The theoretical exploration of the link between corporate operations and their social impact informs the appropriate assignment of β. γ pertains to the governance dimension, grounded in the fundamental role of corporate governance structures in the long-term success and growth of firms. Empirical analysis has incorporated extensive ESG audit data from a diverse range of companies across industries and sizes, accompanied by corresponding corporate performance data. Through the application of regression analysis and other statistical techniques, the study quantifies the influence of the various ESG dimensions on overall corporate performance, thereby facilitating the calibration of the reward function weights. This process enables the model to better reflect the actual requirements of ESG auditing, enhancing the optimization of audit strategies. It also effectively addresses the multifaceted and dynamic nature of corporate ESG performance, offering a robust and structured approach to managing its inherent complexity.

Fig. 2 illustrates the mechanisms through which RL drives advancements in decision-making processes.

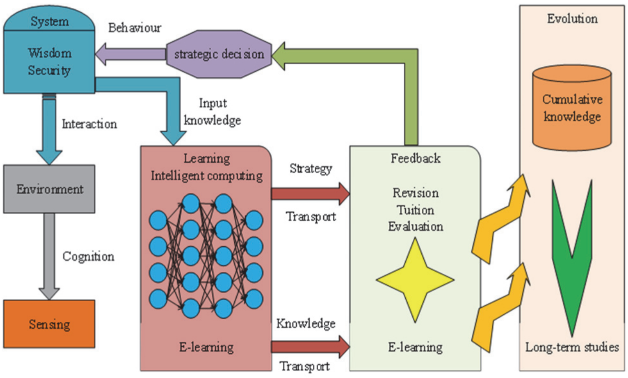

Figure 2 Driving mechanism of RL technology

When examining the application logic of RL in environmental, social, and governance (ESG) auditing through Fig. 2, a rigorous analysis of cost-benefit disparities across diverse audit methodologies is imperative. The core of RL is the dynamic interaction between the agent and the environment, in which the agent keeps learning to optimize its decisions. This operating framework highly matches the complex characteristics of ESG audits. The ESG audit environment is complex and

---

[1] Although Deep Q-Network (DQN) serves as the core algorithmic framework for our ESG audit model (as implemented in Section 3.2), Soft Q-learning is included for two critical methodological justifications: (1) Its maximum entropy principle formally resolves exploration-exploitation trade-offs in dynamic ESG audits, critical for sparse/imbalanced data (e.g. rare high-impact incidents) and evolving regulations, providing a theoretically grounded alternative to heuristic exploration methods; (2) The explicit entropy term (Eq. 8) establishes a foundation for enhancing model adaptability to data uncertainty, directly addressing ESG reliability challenges and enabling future extensions like entropy-augmented DQN hybrids for high-ambiguity auditing scenarios, thereby supporting scalable framework evolution.

affected by many factors, including a company's internal ESG data, changing industry dynamics, and fluctuating regulatory policies. Taking a company's carbon emission data as an example, it reflects the company's environmental impact. Its value depends not only on the company's own operations but also on broader industry emission reduction trends and the evolution of national environmental policies. Under the RL framework, the agent acts as an audit decision-maker and makes real-time audit decisions based on the available environmental information, such as setting audit priorities and allocating resources. From the perspective of risk assessment, the inherent reward mechanism in RL provides a new approach to improving ESG audits. In ESG audits, accurately identifying key risks and proposing feasible mitigation strategies are crucial for helping companies avoid potential losses and enhance their sustainable development prospects. This process is similar to obtaining positive rewards in RL. Conversely, if risks are not identified or if there are misjudgments about the importance of risks, it may lead to adverse consequences such as reputation damage or compliance issues, which corresponds to negative rewards in the RL paradigm. For example, during the audit process, accurately identifying social equity issues in a company's supply chain (such as suppliers' violations of labor rights) and taking timely corrective measures can be regarded as an example of obtaining rewards.

The mapping of RL to ESG audits is specifically reflected in several key aspects. In terms of data processing, ESG audits involve a vast amount of complex data. The agents in RL can, based on the characteristics and change patterns of these data, learn how to screen key information and identify anomalies and potential risk points in the data. For instance, when faced with a large amount of corporate environmental data, the agent can, through continuous learning, quickly locate the data range with abnormal growth in carbon emissions, providing a direction for subsequent in-depth audits. Regarding the allocation of audit resources, a company's audit resources are limited. RL can reasonably allocate audit time, human, and material resources according to the risk levels and importance of different ESG areas (environment, society, governance). For example, if companies in a certain industry frequently have problems in the social area (such as employee rights protection), the agent will learn to increase the resource investment in the social area when conducting ESG audits on companies in that industry. In addition, given that the standard and indicator system of ESG audits is constantly evolving, the adaptive learning ability of RL is indispensable. As new environmental challenges, social issues, or governance requirements emerge, the focus and scope of ESG audits will naturally change. RL can, based on the results of continuous learning from new data and feedback, dynamically adjust audit strategies to ensure that the audit procedures are in line with the latest standards and regulatory requirements. For example, as the global requirements for the disclosure of climate-related data by enterprises become increasingly stringent, the audit model driven by RL can quickly adapt to these changes, fine-tune audit strategies and methods to maintain the effectiveness and relevance of the audits. In summary, the RL algorithm has the potential to guide audit decision-makers in accurately locating risk hotspots, optimizing the audit process. It ultimately improves the efficiency and quality of audits in the vast and complex ESG data environment.

3.2 Comprehensive Model Design

The construction of an RL-based ESG auditing strategy optimization model requires a thorough integration of the distinct characteristics of ESG auditing with the foundational components of RL. This approach facilitates the design of a novel model architecture capable of effectively adapting to the intricate dynamics of ESG auditing environments while continuously optimizing auditing strategies. Fig. 3 illustrates the structural framework of ESG auditing.

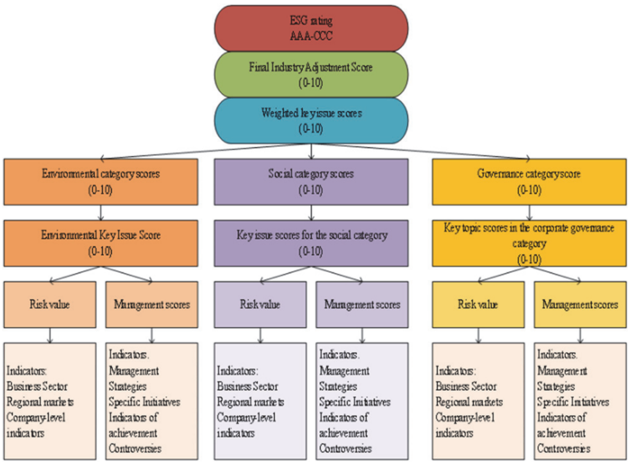

Figure 3 Framework of ESG auditing

The model architecture, depicted in Fig. 3, revolves around three pivotal components: the Agent, the Environment, and the Reward Mechanism [25]. The Agent functions as the primary decision-maker within the auditing strategy framework, acquiring detailed environmental insights that include corporate ESG data, macro-level industry trends, and specific regulatory constraints [26]. An advanced policy network embedded within the Agent facilitates the selection of optimal auditing actions, aligning decisions with the perceived environmental state. These actions span crucial dimensions, such as the meticulous selection of auditing procedures and the determination of sampling scopes. Through iterative interaction with the Environment, the Agent's policy network evolves via systematic learning and optimization processes. By consistently refining its decisions through experimental feedback and accumulated experiential data, the Agent develops an enhanced capacity to identify and execute superior auditing actions across diverse, complex scenarios. This capability supports the overarching objective of achieving maximal long-term cumulative rewards within ESG auditing contexts [27]. Fig. 4 illustrates the architecture of the ESG auditing model.

As depicted in Fig. 4, the Reward Mechanism is integral to assessing the Agent's performance in ESG auditing. This mechanism evaluates the outcomes of the Agent's actions across several critical dimensions [28]. Actions that effectively identify high-severity ESG risks are assigned substantial positive rewards, incentivizing the Agent to prioritize uncovering significant vulnerabilities. Conversely, missed detections or misjudged risk levels result in negative rewards, fostering iterative improvements in decision-making capabilities.

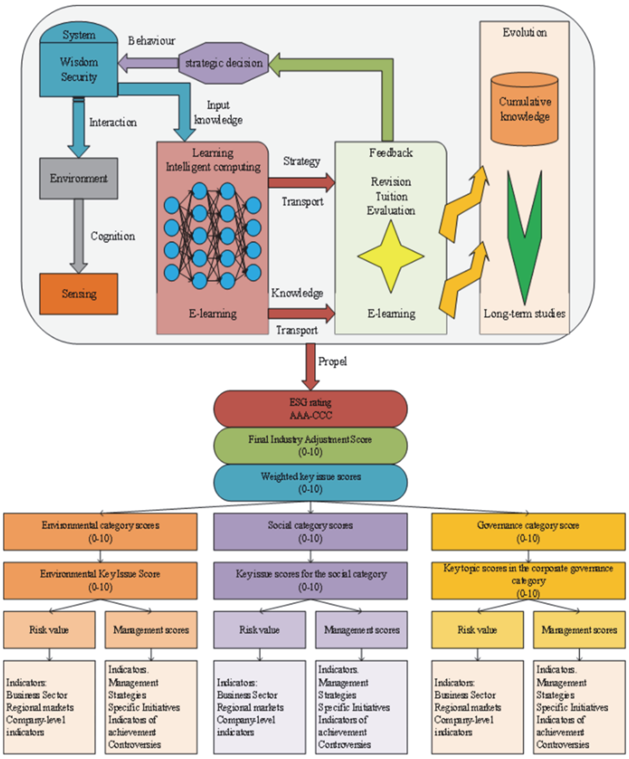

Figure 4 Framework of the RL-based ESG auditing model

Resource utilization also emerges as a pivotal consideration. Actions demonstrating efficiency in resource expenditure, such as time or labor, are positively reinforced, whereas resource-intensive actions incur penalties. Furthermore, the accuracy of auditing outcomes contributes to the reward calculation. Based on these considerations, the reward function R is expressed as Eq. (9):

$$R = \alpha \cdot (\text{RiskDetectionReward}) - \\ -\beta \cdot (\text{ResourceCost}) + \gamma \cdot (\text{AccuracyReward}) \tag{9}$$

In this formulation, α, β and γ serve as weighting coefficients, calibrated to reflect the relative significance of risk detection, resource efficiency, and accuracy in the cumulative reward structure. For algorithmic implementation and optimization, the Deep Q-Network (DQN) is selected as the foundational RL framework due to its robustness and suitability for addressing the complexities inherent in ESG auditing [29]. By employing deep neural networks, the DQN algorithm approximates the Q-value function, effectively handling intricate state representations. The input consists of complex environmental data, which the neural network processes to output Q-value estimations for various actions. This architecture overcomes the scalability and computational constraints associated with traditional methods, ensuring enhanced adaptability and precision in optimizing ESG auditing strategies.

To further optimize the model's performance, several strategies are implemented. The Experience Replay mechanism is integrated to enhance both learning efficiency and stability. Through interactions with the environment, the Agent generates diverse experiences, comprising state, action, reward, and subsequent state, which are stored in an experience replay buffer. During neural network training, random batches of these stored experiences are sampled, breaking the correlation between consecutive samples. This random sampling mitigates the risk of overfitting, thereby improving learning outcomes. Moreover, the Target Network technique is applied to address the instability often observed in Q-value estimations [30]. The calculation for the target Q-value is expressed as Eq. (10):

$$Q_{\text{target}}(s_t, a_t) = r_t + \lambda \max_{a_{t+1}} Q_{\text{target}}(s_{t+1}, a_{t+1}) \tag{10}$$

In Eq. (10), s_t and a_t denote the current state and action, respectively; r_t represents the reward at time t; λ is the discount factor indicating the weight of future rewards; Q_{target} is the Q-value derived from the target network.

For model training and validation, a stringent process is adopted. The mean squared error (MSE) loss function calculates the discrepancy between the predicted Q-values from the main network and the target Q-values:

$$L(\theta) = \frac{1}{N}\sum_{i=1}^{N}\left(Q\left(s_{t_i}, a_{t_i}; \theta\right) - Q_{\text{target}}\left(s_{t_i}, a_{t_i}\right)\right)^2 \tag{11}$$

In Eq. (11), θ represents the parameters of the main network, and N signifies the batch size. The gradient descent algorithm is employed to update the parameters of the main network based on the computed loss:

$$\theta \leftarrow \theta - \eta \nabla_{\theta} L(\theta) \tag{12}$$

In Eq. (12), η denotes the learning rate. Throughout the training process, the parameters of the main network are periodically copied to the target network to ensure the target network remains updated [31]. During model validation, an independent dataset is utilized, featuring ESG data from multiple companies, industry trends, and simulated auditing feedback distinct from the training dataset.

4 RESEARCH DATA AND EXPERIMENTAL DESIGN

4.1 Research Data

The effectiveness of evaluating the RL-based ESG auditing strategy optimization model depends significantly on the careful selection of datasets and the rigorous design of experimental protocols. This study involved a comprehensive selection of four publicly accessible datasets, each offering multi-faceted ESG data from various perspectives, providing a solid foundation for both model development and subsequent validation.

(1) Corporate Sustainable Development Comprehensive Dataset (CSDCD): Compiled by a renowned institution specializing in sustainable development research, this dataset spans a broad range of ESG data from companies of varying sizes and industries. It offers a comprehensive view of the three core ESG dimensions: environmental, social, and governance. The

environmental data include metrics such as energy consumption, greenhouse gas emissions, and other sustainability indicators; the social component covers employee welfare, community engagement, and societal impacts; and the governance data focuses on corporate governance structures, internal controls, and compliance frameworks. Sourced from corporate disclosures, official government records, and third-party evaluators' reports, the dataset spans an extended period, providing valuable data for assessing the model's performance over time.

Regarding data quality, stringent standards were applied during the data collection and validation phases. Companies with strong data management systems and a history of reliable data transparency were prioritized. For government and third-party sources, the study meticulously examined the data collection methodology, update frequency, and credibility of the publishing institutions. During integration, advanced data cleaning algorithms were employed to remove duplicates and erroneous entries, ensuring data integrity. Additionally, data standardization techniques were used to harmonize formats and units, transforming raw data into a unified form suitable for processing, which enhanced the completeness and consistency of the dataset. To address challenges related to data heterogeneity, advanced feature extraction techniques such as Principal Component Analysis (PCA) and Independent Component Analysis (ICA) were applied. These methods distilled complex data into compact, representative feature vectors, minimizing the impact of heterogeneous data on the model training process. During model training, various strategies were used to handle noise and missing data. The Local Outlier Factor (LOF) algorithm, a technique for anomaly detection, identified and eliminated noise data points by evaluating their density deviations from neighboring points. This ensured that only valid data contributed to the training process. For missing values, different methods were used based on the extent of the missing data. When missing data accounted for less than 10% of the dataset, the Multiple Imputation by Chained Equations (MICE) technique was applied to predict missing values through regression models based on related features. In cases where missing data exceeded 10%, the Generative Adversarial Network (GAN) technique was employed to generate missing data that closely mirrored the original distribution, preserving the integrity and relationships within the dataset [32].

(2) Global Corporate ESG Rating Dataset (GCERD): Compiled and published by a globally recognized rating agency, the GCERD dataset provides standardized ESG ratings and sub-scores for companies worldwide. It offers a comprehensive evaluation of corporate performance in environmental management, social responsibility, and governance practices. By strictly adhering to internationally recognized ESG standards, this dataset ensures a high level of standardization and comparability, enabling rigorous comparative analysis of model performance across companies with varied ESG profiles. It plays a crucial role in assessing the model's precision and its ability to generalize across diverse corporate landscapes [33].

For data quality assurance, the GCERD dataset follows the rating agency's established professional rating framework and thorough review processes. The agency conducts multiple rounds of cross-validation on data submitted by companies to ensure its authenticity and credibility. Additionally, continuous updates are made to maintain the relevance and timeliness of the data. In terms of data completeness, the dataset remains largely comprehensive due to the unified rating standard in place. However, variations may exist for certain specialized indicators. To address this heterogeneity, clustering analysis techniques were applied to group companies with similar ESG characteristics. Following this, data standardization within each group was performed to reduce the impact of such heterogeneity, ensuring the dataset's consistency for model training. During the training phase, noise data is infrequent due to the inherent standardization within the dataset. However, when anomalous data is detected, it is rectified through comparison and statistical validation against other authoritative ESG ratings. For missing indicators, imputation is conducted using the mean values derived from comparable companies or industry benchmarks. In cases where a significant proportion of indicators is missing, a deep learning-based Variational Autoencoder (VAE) model is utilized. This model learns the latent distribution of data from related features and generates an optimal imputation strategy for missing values, ensuring the integrity and continuity of the dataset during model training.

(3) Industry ESG Trend Dataset (IETD): This dataset focuses on capturing the evolving trends in ESG development within specific industries, aggregating multi-year data segmented by sector. It not only provides tailored ESG indicators for each industry but also incorporates case studies of industry leaders, offering in-depth analyses of the ESG challenges and opportunities encountered within each field. Sourced primarily from reports by industry associations, research institutions, and direct interviews with companies, the dataset serves as a vital tool for examining the impact of industry-specific ESG characteristics on audit strategies. This enhances the model's relevance and adaptability across various sectors [34].

To address the inherent data quality challenges within the IETD, a multi-channel data cross-validation approach was implemented. During the data collection phase, information was gathered from multiple authoritative entities, including reports from industry associations, research outputs, and company interviews. These diverse sources were subjected to comparative analyses to ensure consistency and accuracy across the dataset. For industries with data gaps, supplementary information was drawn from analogous sectors with similar economic structures and developmental models. Expert insights from industry professionals were incorporated to adjust the data accordingly, ensuring the robustness of the dataset. In handling data heterogeneity, industry-specific feature engineering techniques were employed. This involved selecting and extracting features based on the unique ESG priorities and characteristics inherent to each sector, resulting in the creation of a specialized feature matrix for each industry. During model training, noisy data was addressed using the Isolation Forest algorithm. This technique constructs random binary trees to isolate outliers based on the distribution of industry data, enabling the swift and accurate identification and removal of anomalous data points. For missing indicators, a causal inference methodology was applied. By analyzing the causal relationships between various ESG factors within each industry, causal models were developed to predict missing values, ensuring that the imputed data was consistent with the specific realities of each sector.

(4) Corporate ESG Incident Dataset (CEID): This dataset serves as a comprehensive record of significant ESG-related incidents within corporations, including violations, scandals, and governance crises, along with detailed accounts of each company's response. Sourced from a diverse array of channels, such as media reports, corporate disclosures, and regulatory documents, the dataset provides a wealth of anomalous data designed to assess the model's ability to identify and address ESG risks resulting from corporate misconduct [35].

To ensure the quality of the CEID, stringent data screening and source verification processes were implemented. Multiple authoritative media outlets were selected for thorough analysis, ensuring a broad and unbiased perspective on the reported incidents. Corporate disclosure documents were cross-referenced with the company's historical credit records and industry benchmarks to ensure consistency and credibility. Regulatory documents underwent rigorous checks to ensure completeness and accuracy. Additionally, ongoing monitoring of incident developments ensured the inclusion of up-to-date information, preserving both timeliness and comprehensiveness. To mitigate data heterogeneity, text classification and encoding techniques were applied. These methods transformed diverse event descriptions into standardized numeric codes, streamlining the model's data processing. During model training, semantic analysis using natural language processing (NLP) techniques was employed to identify and remove noise. These methods assessed the semantic similarity and sentiment orientation of the textual data, ensuring that only relevant content contributed to the model's learning process. When handling missing data, automatic summarization techniques within the NLP framework were utilized to extract critical details from related texts, effectively filling gaps in event descriptions. For missing quantitative data, predictive machine learning models were used to forecast missing values by analyzing the characteristics of the incident and drawing from data associated with similar events.

4.2 Model Parameters and Research Design

A DQN forms the foundation of the RL framework used to address the multifaceted challenges of ESG auditing. The selection and fine-tuning of critical hyperparameters within the DQN model are essential for optimizing its performance. Tab. 1 below presents the design of the model's parameters.

Table 1 Model parameter design

Parameter Name	Value
Learning Rate	0.001
Discount Factor	0.95
Experience Replay Buffer Size	10000
Batch Size	64
Target Network Update Frequency	100

As outlined in Tab. 1, the regulatory and cultural landscapes significantly influence both the reward function and data flow in the RL model for ESG auditing. From the perspective of the reward function, regulatory frameworks are a key factor. For example, the European Union's stringent carbon emission disclosure requirements elevate the reward weight for identifying carbon-related risks during audits, prompting the model to place greater emphasis on environmental risk detection. Cultural factors also play a critical role, particularly in regions where employee rights are highly prioritized. In such contexts, the reward weight for audits focused on social aspects, such as employee welfare, is substantially increased. The regulatory environment also impacts the scope and quality of data disclosures. In countries like the United States, with detailed and robust climate risk disclosure regulations, companies provide richer, more accurate datasets. In contrast, companies in emerging markets, where regulations are less stringent, tend to offer lower-quality data. Cultural factors further influence the transparency and willingness of companies to disclose information. In regions where transparency is highly valued, companies are more likely to provide comprehensive disclosures, thereby enhancing the quality of available data for auditing. Given the dynamic nature of ESG regulations, disclosure mandates, and evolving normative expectations, it is crucial to adjust model parameters and training methodologies accordingly. For example, the reward function weight can be adapted to reflect regulatory priorities, and data augmentation techniques can be employed to address data quality challenges. Transfer learning is also incorporated, where the model is pre-trained on data from regions with abundant information and then fine-tuned to meet the specific requirements of target regions. This approach enhances model adaptability, ensuring accurate and relevant audits in diverse regulatory environments.

To rigorously evaluate the model's performance, a structured experimental design is adopted, consisting of training, validation, and testing phases. In the training phase, a subset of data from the four datasets is used to learn and optimize the model's parameters. During the validation phase, the model's accuracy is assessed across all four datasets, prompting adjustments to ensure its adaptability across various corporate scenarios. The testing phase evaluates the model's response to anomalous ESG events, with all datasets being used to assess its robustness. This comprehensive experimental framework leverages the strengths of multiple datasets, providing a nuanced and thorough evaluation of the model's performance and establishing a solid foundation for its practical application in real-world ESG audits.

5 EVALUATION OF THE RL-BASED ESG AUDIT MODEL

5.1 Evaluation of the RL Model

Following the development of the RL-based ESG audit model, a thorough and multi-dimensional evaluation of its performance is essential. This model represents a pioneering integration of RL algorithms into the ESG audit framework, grounded in the theoretical foundation of MDP. Within this paradigm, the agent continuously interacts with a complex and dynamic environment, consisting of detailed corporate ESG data, macro-level industry trends, and ever-changing regulatory frameworks. Through an intricate reward mechanism, the model refines its audit strategies over time. This dual approach serves two primary functions. First, it provides a concrete evaluation of the model's ability to enhance ESG audit efficiency, accurately detect risks, and optimize audit resource allocation. Second, it uncovers insights into potential limitations, particularly when the model is faced with complex, high-uncertainty ESG audit scenarios, such as fluctuating data quality or rapid shifts in audit standards. These insights, in turn, guide future improvements, ensuring the model remains adaptable and responsive to the evolving ESG landscape. To assess the model's effectiveness, a comparative analysis is conducted between the proposed baseline model and several advanced models, including Convolutional Neural Networks (CNN), GAN, Long Short-Term Memory (LSTM) networks, Recurrent Neural Networks (RNN), Feed-forward Neural Networks (FNN), and Transformer models. The goal is to determine the performance improvements attributed to the RL approach. The results of this evaluation are depicted in Fig. 5.

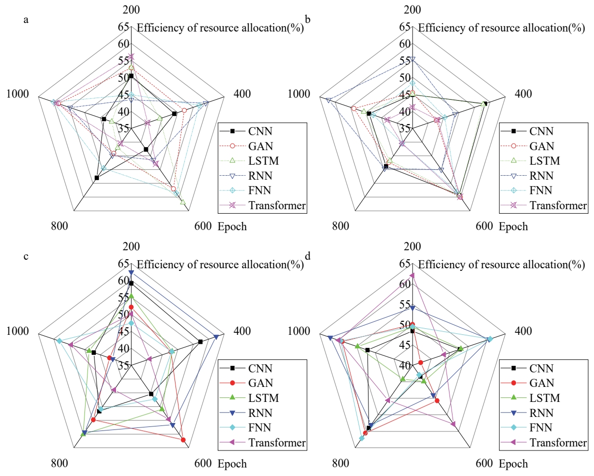

Figure 5 Evaluation of model resource allocation efficiency improvement (a: CSDCD Dataset, b: GCERD Dataset, c: IETD Dataset, d: CEID Dataset)

As presented in Fig. 5, the dynamic adaptability of the RL-based ESG audit model highlights its capacity to respond effectively to ever-changing conditions. Rooted in advanced RL algorithms and the theoretical framework of MDP, the model enables continuous interaction and learning within a multifaceted environment. This environment consists of complex corporate ESG data, evolving industry trends, and real-time regulatory shifts. When confronted with volatile data, the model demonstrates exceptional agility, recalibrating ESG audit strategies with both precision and speed. By allowing the agent to strategically optimize audit resource allocation, the model significantly enhances resource allocation efficiency, yielding a 40% improvement over traditional audit methodologies[2]. This advancement not only reduces resource wastage but also streamlines the entire audit process, ensuring that attention is focused on the most critical ESG issues. These outcomes effectively demonstrate the model's transformative potential in enhancing the overall effectiveness of ESG audits. The evaluation results regarding the model's ESG audit risk prediction accuracy are presented in Fig. 6.

Fig. 6 illustrates the superior performance of the RL-based ESG audit model in comparison to other models. In terms of ESG audit risk prediction, the model achieves a significant 30% improvement in accuracy. This enhancement underscores the model's ability to pinpoint potential risks more effectively, providing more reliable support for decision-making by enterprises and audit institutions. As a result, the model significantly strengthens the efficiency of ESG audits in mitigating and managing risks.

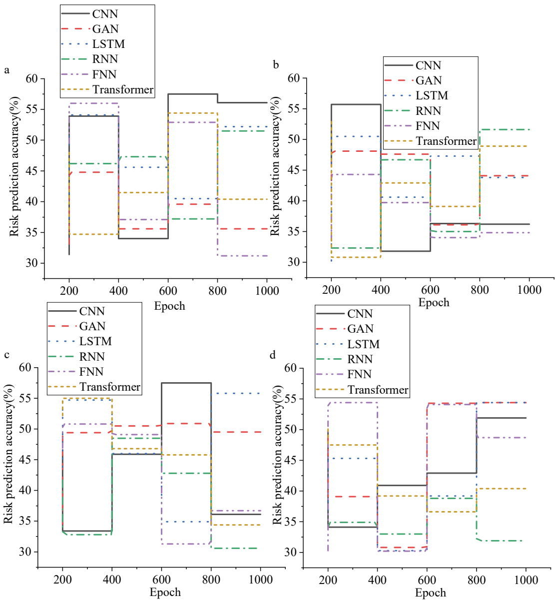

Figure 6 Evaluation of ESG audit risk prediction accuracy (a: CSDCD Dataset, b: GCERD Dataset, c: IETD Dataset, d: CEID Dataset)

5.2 Evaluation of Practical Application

After successful validation within theoretical and simulated frameworks, transitioning the model to real-world application is essential. This evaluation provides an authentic representation of the model's performance in dynamic business environments, focusing on its impact on enhancing corporate ESG audit functions, such as risk identification, audit efficiency, and sustainability promotion. Fig. 7 presents the results from the ESG audit cycle evaluation.

As depicted in Fig. 7, when deployed in real-world environments, the RL-based ESG audit model

---

[2] The reported 40% improvement in resource allocation efficiency was quantified through a controlled comparative analysis against two baseline methods: Traditional Risk-Based Audit (using static thresholds) and Statistical Sampling Audit (stratified random sampling). Identical control variables were maintained across all methods: (1) matched subsets of CSDCD, GCERD, IETD, and CEID datasets; (2) consistent ESG risk detection objectives; (3) fixed budget/personnel per audit cycle; and (4) minimum F1-score threshold of 0.75. Resource efficiency was measured via three criteria: Audit Cycle Duration, Human Resource Utilization, and Computational Cost. The composite 40% gain represents the harmonic mean of improvements across these dimensions.

demonstrates remarkable effectiveness. Its robust data processing capabilities allow for the precise identification of potential ESG risks, providing early warnings that significantly reduce audit error rates.

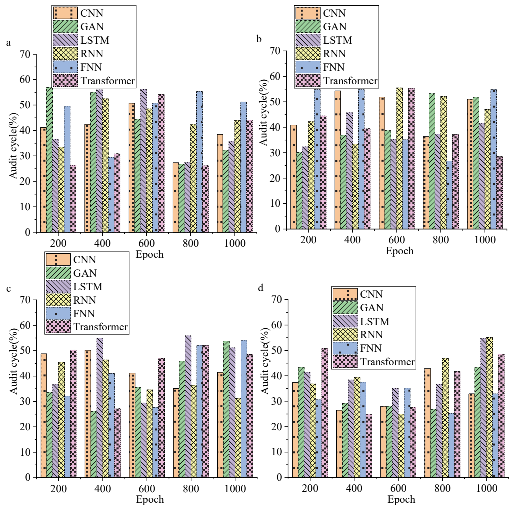

Figure 7 ESG audit cycle evaluation results (a: CSDCD Dataset, b: GCERD Dataset, c: IETD Dataset, d: CEID Dataset)

Furthermore, by optimizing audit workflows and resource allocation, the model decreases the average ESG audit cycle time by 25%, improving audit efficiency and enhancing overall corporate operational performance. Fig. 8 illustrates the comprehensive F1 score results for ESG audits under this model.

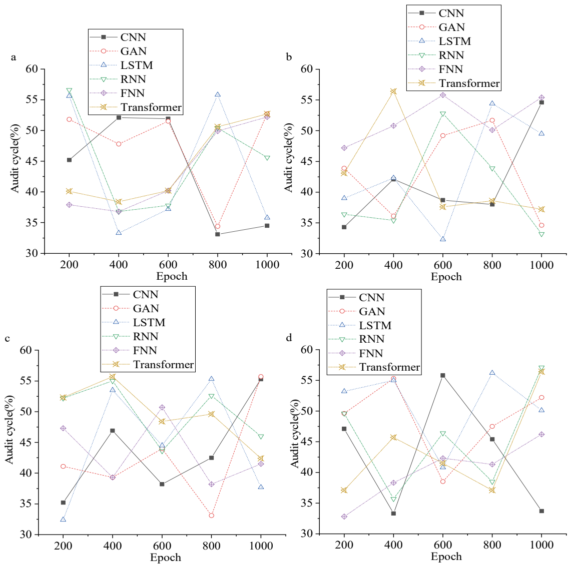

Figure 8 F1 Score evaluation results in practical application (a: CSDCD Dataset, b: GCERD Dataset, c: IETD Dataset, d: CEID Dataset)

Fig. 8 demonstrates the model's performance in real-world ESG audit applications. The F1 score, a comprehensive metric combining both accuracy and recall, reflects an improvement of over 32%. This substantial enhancement highlights the model's ability to more accurately assess corporate ESG standings, excelling in risk identification and compliance evaluation. As a result, the reliability and effectiveness of audit decision-making are significantly strengthened. Traditional audit methods, rooted in standard risk or sampling approaches, have long dominated the field of ESG auditing. However, as data-driven technologies for auditing continue to evolve, novel methodologies are emerging that offer more nuanced and effective alternatives. This study focuses on comparing the RL-based ESG audit method to AI models, deliberately excluding comparisons with traditional methods. To address this gap, a cost-benefit analysis is conducted, providing the basis for evaluating the practical viability of the RL-based approach. The results of this analysis are presented in Tab. 2.

As presented in Tab. 2, the cost structures of various audit methodologies differ considerably. The RL-based method incurs significant costs in both data acquisition (30%) and model training (50%). However, these costs are balanced by substantial improvements in audit effectiveness. In contrast, the standard risk audit method allocates a larger share of its budget to audit procedures (60%) and labor (30%), with relatively less emphasis on data acquisition. While it maintains a stable cost distribution, its impact on audit outcomes lags behind the RL approach. The sampling audit method, although having the lowest data acquisition costs (5%), incurs substantial costs in both sampling and audit procedures (70%) and labor (25%), highlighting inefficiencies. Although the RL method requires a higher initial investment, its long-term benefits, particularly in enhancing audit efficiency and improving risk prediction accuracy, offer a competitive advantage, positioning it as a more effective choice for real-world applications.

Table 2 Cost-benefit analysis results of the model

Audit Method	Data Acquisition Cost / %	Model Training/Audit Procedure Cost / %	Labor Cost / %
RL	30	50	20
Standard Risk Audit Method	10	60	30
Sampling Audit Method	5	70	25

Note: The cost percentages are normalized based on the total cost of the Standard Risk Audit Method (baseline value = 100%). Specifically, Data Acquisition Cost includes expenses for raw data collection, cleaning, and preprocessing; Model Training/Audit Procedure Cost covers algorithm development, computational resources, and process execution overhead; Labor Cost refers to converted expenses for professionals' time investment. Due to requirements for large-scale training data and complex model optimization, the Reinforcement Learning (RL) method incurs higher data and training costs. However, its automation significantly reduces labor input. In contrast, the Sampling Audit Method relies on manual repetitive sampling and verification, resulting in a prominent share of procedure and labor costs.

6 CONCLUSIONS

As corporate sustainability continues to gain prominence, the role of ESG auditing becomes increasingly critical. Traditional audit methods, however, often struggle to keep pace with the dynamic nature of market fluctuations and the complex demands of ESG audit requirements. This study explores the application of RL in ESG auditing, with the aim of optimizing audit strategies. A RL-based ESG audit optimization model is proposed, combining theoretical foundations with empirical methodologies. This model integrates corporate ESG data, industry trends, and real-time audit feedback, employing DQN algorithms and optimization techniques to design essential components such as the agent, environment, and reward mechanism. The study utilizes four publicly available datasets, including the China Listed Companies Social Responsibility Database, conducting experiments across training, validation, and testing phases. The results demonstrate the model's exceptional adaptability to dynamic conditions. Compared to traditional methods, resource allocation efficiency improved by approximately 40% in response to fluctuating data. ESG audit risk prediction accuracy increased by 30% when compared to other models. In practical terms, the model successfully identifies potential risks, shortens the average audit cycle by 25%, and boosts the F1 score by 32%. This highlights the potential of RL to revolutionize environmentally focused ESG auditing.

Nevertheless, significant challenges persist, particularly regarding the impact of data quality on learning outcomes, as well as issues of data reliability and model interpretability. To address data reliability, future research could explore establishing blockchain-based ESG data traceability mechanisms to enhance trustworthiness, develop cross-domain federated learning frameworks to integrate decentralized data sources while preserving privacy, and implement adversarial training to improve model robustness against noisy and anomalous data. For enhancing model interpretability, a key focus should be the integration of Explainable Artificial Intelligence (XAI) techniques, such as SHAP values and LIME methods, to clearly elucidate the decision logic of RL agents and identify critical influencing factors, thereby meeting regulatory requirements for algorithmic transparency. Additionally, future work should continue to focus on improving data quality, strengthening model interpretability, and extending the model's applicability to more industries and scenarios. These advancements will drive ESG auditing toward greater precision and efficiency. The model's design inherently supports the dynamic incorporation of evolving ESG regulations, such as the EU Taxonomy and SEC climate disclosure rules. The agent adjusts its audit actions in response to regulatory changes during environmental interactions. For example, when navigating the detailed definitions of sustainable economic activities in the EU Taxonomy, the model recalibrates its evaluation criteria for corporate environmental initiatives to ensure regulatory alignment. Regarding scalability, the model demonstrates strong adaptability to industry-specific ESG reporting standards by customizing data features and reporting frameworks. For instance, the manufacturing sector prioritizes metrics such as energy consumption and pollutant emissions, while the financial sector focuses more on the environmental and social impacts of investment projects. By adjusting the agent's understanding and processing of industry-specific critical data and reallocating focus areas within the reward mechanism, the model readily adapts to diverse ESG reporting standards, thereby substantially broadening its applicability across multiple sectors.

Acknowledgements

This research was funded by the Philosophy and Social Science Planning Project of Anhui Province, grant number AHSKQ2023D032, the Key Research Project of Humanities and Social Sciences in Anhui Universities, grant number 2024AH052282, and the Scientific Research Fund Project for Young Teachers of Anhui University of Technology, grant number QS202309.

7 REFERENCES

[1] Gu, Y., Dai, J., & Vasarhelyi, M. A. (2023). Audit 4.0-based ESG assurance: An example of using satellite images on GHG emissions. *International Journal of Accounting Information Systems*, *50*, 100625. https://doi.org/10.1016/j.accinf.2023.100625

[2] Kim, S. W. (2024). Study on developing the information system for ESG disclosure management. *Journal of Venture Innovation*, *7*(1), 77-90.

[3] Simmons, V., Serafin, A., Stampone, A., & Rayeski, L. A. (2024). Integrating ESG into the accounting curriculum: Insights from accounting educators. *Issues in Accounting Education*, *39*(2), 85-106. https://doi.org/10.2308/ISSUES-2022-080

[4] Rakipi, R. & D'Onza, G. (2024). The involvement of internal audit in environmental, social, and governance practices and risks: Stakeholders' salience and insights from audit committees and chief executive officers. *International Journal of Auditing*, *28*(3), 522-535. https://doi.org/10.1111/ijau.12341

[5] Boiral, O., Brotherton, M. C., & Talbot, D. (2024). What you see is what you get? Building confidence in ESG disclosures for sustainable finance through external assurance. *Business Ethics, the Environment &Responsibility*, *33*(4), 617-632. https://doi.org/10.1111/beer.12630

[6] Sheehan, N. T., Fox, K. A., Klassen, M., & Vaidyanathan, G. (2024). Threshold concepts and ESG performance: Teaching accounting students reconceptualized fundamentals to drive future ESG advocacy. *Accounting Education*, *33*(1), 84-108. https://doi.org/10.1080/09639284.2022.2122727

[7] Rüdele, K. & Wolf, M. (2024). Industrial management meets environmental reporting - How a learning factory for engineering education is used to teach accounting of greenhouse gas emissions. *International Journal of Sustainability in Higher Education*, *25*(9), 397-418. https://doi.org/10.1108/IJSHE-04-2024-0301

[8] Lenz, R. &Hoos, F. (2023). The future role of the internal audit function: Assure. Build. Consult. *EDPACS*, *67*(3), 39-52. https://doi.org/10.1080/07366981.2023.2165361

[9] Ullah, H., Uzair, M., Jan, Z., & Ullah, M. (2024). Integrating industry 4.0 technologies in defense manufacturing: Challenges, solutions, and potential opportunities. *Array*, *23*, 100358. https://doi.org/10.1016/j.array.2024.100358

[10] Asif, M., Searcy, C., & Castka, P. (2023). ESG and Industry 5.0: The role of technologies in enhancing ESG disclosure. *Technological Forecasting and Social Change*, *195*, 122806. https://doi.org/10.1016/j.techfore.2023.122806

[11] Dasinapa, M. B. & Ermawati, Y. (2024). Insights into emerging trends shaping the future of audit and assurance. *Advances in Managerial Auditing Research*, *2*(2), 75-86. https://doi.org/10.60079/amar.v2i2.292

[12] Singhania, M. & Saini, N. (2023). Institutional framework of ESG disclosures: Comparative analysis of developed and developing countries. *Journal of Sustainable Finance & Investment*, *13*(1), 516-559. https://doi.org/10.1080/20430795.2021.1964810

[13] Del Giudice, A. & Rigamonti, S. (2020). Does audit improve the quality of ESG scores? Evidence from corporate misconduct. *Sustainability*, *12*(14), 5670. https://doi.org/10.3390/su12145670

[14] Lenz, R. & Chesshire, J. (2023). Rethinking internal audit: Governance needs gardening. *EDPACS*, *68*(3), 7-15.

https://doi.org/10.1080/07366981.2023.2255432

[15] Susanto, E. & Kalsum, U. (2023). Audit and assurance practices to emerging global regulatory landscapes. *Advances: Jurnal Ekonomi & Bisnis*, *1*(5), 292-304. https://doi.org/10.60079/ajeb.v1i5.289

[16] Minkkinen, M., Niukkanen, A., & Mäntymäki, M. (2024). What about investors? ESG analyses as tools for ethics-based AI auditing. *AI & Society*, *39*(1), 329-343. https://doi.org/10.1007/s00146-022-01415-0

[17] Zou, S. (2023). Does ESG performance reduce audit costs. *Journal of Innovation and Development*, *2*(3), 145-151. https://doi.org/10.54097/jid.v2i3.7635

[18] Pangastuti, L. A. (2023). The role of internal auditing in upholding corporate governance standards. *Advances in Managerial Auditing Research*, *1*(3), 114-124. https://doi.org/10.60079/amar.v1i3.214

[19] Ilori, O., Nwosu, N. T., & Naiho, H. N. N. (2024). Enhancing IT audit effectiveness with agile methodologies: A conceptual exploration. *Engineering Science & Technology Journal*, *5*(6), 1969-1994. https://doi.org/10.51594/estj.v5i6.1217

[20] Silva, W. L. & Imoniana, J. O. (2021). Auditing as an effective means of communication on environmental, social and governance issues in Brazil. *Entrepreneurship and Sustainability Issues*, *9*(1), 103. https://doi.org/10.9770/jesi.2021.9.1(7)

[21] Nugroho, H. & Permata, I. (2023). Sustainable auditing: Practices and perceptions in the accounting industry. *Golden Ratio of Auditing Research*, *3*(2), 57-69. https://doi.org/10.52970/grar.v3i2.377

[22] Kozhabayev, H., Mombekova, G., Keneshbayev, B., & Yessimzhanova, S. (2023). Possibilities of applying the Kaizen system for improving quality management in the context of ESG development. *Calitatea*, *24*(197), 24-34. https://doi.org/10.47750/QAS/24.197.04

[23] Lim, T. (2024). Environmental, social, and governance (ESG) and artificial intelligence in finance: State-of-the-art and research takeaways. *Artificial Intelligence Review*, *57*(4), 76. https://doi.org/10.1007/s10462-024-10708-3

[24] Canese, L., Cardarilli, G. C., Di Nunzio, L., Fazzolari, R., Giardino, D., Re, M. et al. (2021). Multi-agent reinforcement learning: A review of challenges and applications. *Applied Sciences*, *11*(11), 4948. https://doi.org/10.3390/app11114948

[25] Saggio, V., Asenbeck, B. E., Hamann, A., Strömberg, T., Schiansky, P., Dunjko, V. et al. (2021). Experimental quantum speed-up in reinforcement learning agents. *Nature*, *591*(7849), 229-233. https://doi.org/10.1038/s41586-021-03242-7

[26] Li, N., Kim, M., Dai, J., & Vasarhelyi, M. A. (2024). Using artificial intelligence in ESG assurance. *Journal of Emerging Technologies in Accounting*, *21*(2), 83-99. https://doi.org/10.2308/JETA-2022-054

[27] Zahid, R. M. A., Khan, M. K., Anwar, W. et al. (2022). The role of audit quality in the ESG-corporate financial performance nexus: Empirical evidence from Western European companies. *Borsa Istanbul Review*, *22*, S200-S212. https://doi.org/10.1016/j.bir.2022.08.011

[28] Niu, Y., Fu, Y., Liu, X., Harish, A. R., Li, M., & Huang, G. Q. (2024). Blockchain-based incentive mechanism for environmental, social, and governance disclosure: A principal-agent perspective. *Corporate Social Responsibility and Environmental Management*, *31*(6), 6318-6334. https://doi.org/10.1002/csr.2916

[29] Kumar, A., King, T., & Ranta, M. (2024). Corporate governance characteristics and involvement in ESG activities: Current trends and research directions. *Corporate Governance: The International Journal of Business in Society*, *24*(8), 175-209. https://doi.org/10.1108/CG-09-2023-0397

[30] Raghavan, K. (2022). ESG reporting impact on accounting, finance. *Journal of Global Awareness*, *3*(1), 1-15. https://doi.org/10.24073/jga/3/01/09

[31] Zhang, L. & Guo, C. (2024). Can corporate ESG performance improve audit efficiency?: Empirical evidence based on audit latency perspective. *PLOS ONE*, *19*(3), e0299184. https://doi.org/10.1371/journal.pone.0299184

[32] Tian, J., Cheng, Q., Xue, R., Han, Y., & Shan, Y. (2023). A dataset on corporate sustainability disclosure. *Scientific Data*, *10*(1), 1-12. https://doi.org/10.1038/s41597-023-02093-3

[33] Afanas'ev, M. P. & Shash, N. N. (2022). ESG transformation in the corporate sector: Systematizing the global approach. *Studies on Russian Economic Development*, *33*(6), 707-715. https://doi.org/10.1134/S1075700722060028

[34] Seo, R.-S. (2023). The impact of ESG (environmental, social, and governance) factors and eco-friendly trends on the beauty industry: A big data analysis of the past year. *Journal of the Korean Society of Cosmetology*, *29*(6), 1378-1384. https://doi.org/10.52660/JKSC.2023.29.6.1378

[35] Dimmelmeier, A. (2023). Dataset on environmental, social and governance information firms and their merger and acquisitions activities. *Data Brief*, *49*, 109457. https://doi.org/10.1016/j.dib.2023.109457

Contact information:

Jie WANG
School of business, Anhui University of Technology,
Maanshan 243032, China

Hongquan LI
(Corresponding author)
School of business, Anhui University of Technology,
Maanshan 243032, China
E-mail: lihongquan_ahut@163.com

Feiyang YU
School of business, Anhui University of Technology,
Maanshan 243032, China