COMPARING SHORT VS. LONG-TERM DECISION STRATEGIES IN SUSTAINABLE HUMAN RESOURCE MANAGEMENT: A DEEP REINFORCEMENT LEARNING APPLICATION

Vahid Gholamzadeh[a], Adel Guitouni[a], and Behrooz Khorshidvand[a]

[a] Peter B. Gustavson School of Business, University of Victoria, Victoria, BC, Canada
{vahidgh, adelg, behroozkhorshidvand}@uvic.ca

ABSTRACT

This study investigates the comparative effectiveness of long-term versus short-term decision-making strategies in human resources management within the agri-food supply chain, focusing specifically on blueberry harvesting in BC, Canada. We apply a *Deep Reinforcement Learning* (DRL) model to assess the performance over a long-time versus short-term planning horizon of decision-making strategies in the context of sustainable blueberry harvesting. We consider the economic and social dimensions of sustainability to evaluate the quality of the decisions made by the farmer agent. The economic farmer agent aims to maximize profitability through increased harvesting productivity, while the social farmer agent seeks to create jobs and reduce job turnover. We model the problem as a sequential dynamic decision problem and apply DRL to train farmer agents with different decision-making personas. The results suggest that not only does the long-term strategy yield greater profitability for farmers, but it also leads to improved social performance.

Keywords: supply chain management, agri-food industry, deep reinforcement learning.

1 INTRODUCTION

In the face of ever-evolving challenges, the agricultural industry remains at the forefront of sustaining global food security and economic stability. Despite significant advancements in productivity and technological innovations, this sector grapples with a myriad of obstacles that threaten the integrity of agri-food supply chains and jeopardize the efficacy of food security programs [1]. Among these challenges, climate change emerges as a critical exacerbating factor, contributing to the destabilization of climate conditions and highlighting the imperative for implementing robust sustainability measures within agriculture [2]. The increasing concentrations of greenhouse gases are a principal driver of climate instability, emphasizing the need for the sector to adapt through sustainable practices in the face of varied climate change scenarios [3].

The decision-making processes within the agri-food supply chain are significantly influenced by the perception of climate variability elements, namely rainfall, temperature, and sunshine. However, a pronounced gap exists between the farmers' perceptions of climate variability and the empirical data provided by scientific research, spotlighting the crucial role of aligning short-term and long-term *Human Resource Management* (HRM) strategies to bridge this divide [4]. Developing a cohesive understanding of these strategies, including integrating sustainability objectives, is paramount in fostering stakeholder consensus toward achieving sustainable yield production that meets consumer demand. Despite the acknowledged impact of climate change on the severity and patterns of extreme weather events [5], more research is still needed to investigate the broader implications of climate variability on food availability, particularly at national and regional scales [6]. This gap underscores the urgency of exploring sustainable agricultural practices that can withstand the challenges posed by climate change, ensuring the resilience of food supply chains against the backdrop of global environmental shifts.

This paper addresses these complex challenges by adopting a holistic view of sustainability within the agricultural sector, guided by the *Triple Bottom Line* (TBL) approach that balances environmental, economic, and social goals [7]. By focusing on sustainable HRM strategies and the potential of novel

Proc. of the 2024 Annual Simulation Conference (ANNSIM'24), May 20-23, 2024, American University, DC, USA
C. Ruiz-Martin, B. Oakes and R. Cárdenas, eds.

decision-making methodologies like *Deep Reinforcement Learning* (DRL), this study aims to enhance traditional agricultural practices and decision-making processes. Through comprehensive analyses and the exploration of multiple scenarios, the paper offers actionable insights and practical solutions to optimize agricultural efficiency, support the sustainability of the agri-food supply chain, and improve informed decision-making within the industry.

The paper is organized as follows: Section 2 reviews relevant literature and goals. Section 3 outlines the problem area. Section 4 explains the model and algorithm. Section 5 shares results. Finally, Section 6 concludes with implications and future research directions.

2 LITERATURE REVIEW

The scientific consensus is that greenhouse gases significantly impact climate and agriculture, potentially reducing crop yields. The IPCC stresses the need for changes in farming and land use to adapt to shifting climate patterns, such as temperature and rainfall. The complexity of climate's effects on crops complicates understanding and adapting to these changes [8].

Studying climate change's impact on the agri-food supply chain requires analyzing climate indicators to produce realistic forecasts. Acknowledging climate change's importance, understanding future scenarios is essential for stakeholders like suppliers, farmers, buyers, end-users, and policymakers [9]. This highlights the importance of accurate decision-making in agriculture. Reliable forecasts can strengthen the agri-food supply chain, improve food security through strategic imports and exports, and help farmers make informed financial and operational choices based on climate projections [10]. Getting accurate crop yield forecasts is tough but vital for agri-food chain goals. Sustainability, especially in the food sector facing complex social and environmental issues, has become crucial. With growing public interest, more companies are embedding sustainability into their strategies [11].

Sustainability means meeting current needs without compromising future generations' ability to meet theirs, encompassing social, environmental, and economic responsibilities. The TBL framework, which balances these objectives, is widely used by companies to assess their supply chain performance [12]. Meeting the TBL for sustainability is challenging for decision-makers due to unpredictable factors like yield expectations, lack of transparency, and the complexity of socioeconomic systems, leading to uncertainty. Environmental issues, often without clear scientific explanations, add to these complications [13].

Researchers suggest viewing decisions as parts of complex, adaptive systems to manage challenges effectively. With the rise in supply chain disruptions affecting various aspects, enhancing resilience is now a top priority. A study shows 80% of decision-makers emphasize resilience to mitigate unforeseen events, indicating a move towards developing stronger coping mechanisms [14, 15]. Our research emphasizes the crucial role of HRM in aligning recruitment with corporate strategies, values, and labor market conditions. It shows that successful recruitment involves not just attracting talent but also integrating them into company culture and enhancing their performance. Choosing between external hires and internal promotions is key to recruitment strategies. Moreover, high turnover can impact skill consistency and customer service, highlighting the need for continuous training for new hires [16]. For lasting success, companies should focus on identifying potential leaders from new hires, aligning recruitment with organizational objectives. This approach, supported by leadership and motivation research, is crucial for sustainable growth [17].

This study contributes significantly to agricultural science by integrating DRL into existing frameworks, thus refining, and enhancing traditional methodologies. By applying DRL, we achieve a nuanced and dynamic evaluation of decision-making strategies, enabling a comprehensive assessment of their performance across various scenarios. This sophisticated analysis equips decision-makers with essential insights into the adaptability of strategies, supporting more informed and strategic agricultural planning. Moreover, transcending theoretical discourse, the research delivers actionable outcomes through detailed scenario analysis, offering indispensable resources to stakeholders in agriculture and decision science. The findings validate the models and strategies developed, providing pragmatic guidance to enhance decision-making processes, optimize agricultural practices, and ensure the sustainability of agri-food supply chains.

3 MODELING AND SIMULATION

This study analyzes sustainable HRM strategies in British Columbia's blueberry supply chain during harvest season. It aims to evaluate the short and long-term impacts of labor hiring choices, such as whether to hire seasonally or establish ongoing contracts. The decisions farmers face regarding agricultural labor has significant economic, social, and environmental ramifications. The study focuses on a blueberry farm divided into blocks, each with unique factors affecting HRM strategy.

The problem statement includes the evaluation of different decision-making strategies over multiple years, focusing on their cumulative impact on sustainability. As shown in Figure 1, each decision-making juncture (S_t, where t ranges from 1 to T_n, T_n representing the entire harvesting season in year *n* over a planning horizon of *N* years) presents the farmer with a complex array of choices regarding the optimal composition of the labour force, contingent upon expected yield outcomes and broader sustainability objectives. The repercussions of these decisions, encapsulated within a reward framework comprising financial and non-financial value functions (e.g., profit, social well-being), serve as the precursors for subsequent strategic adjustments (S_{t+1}).

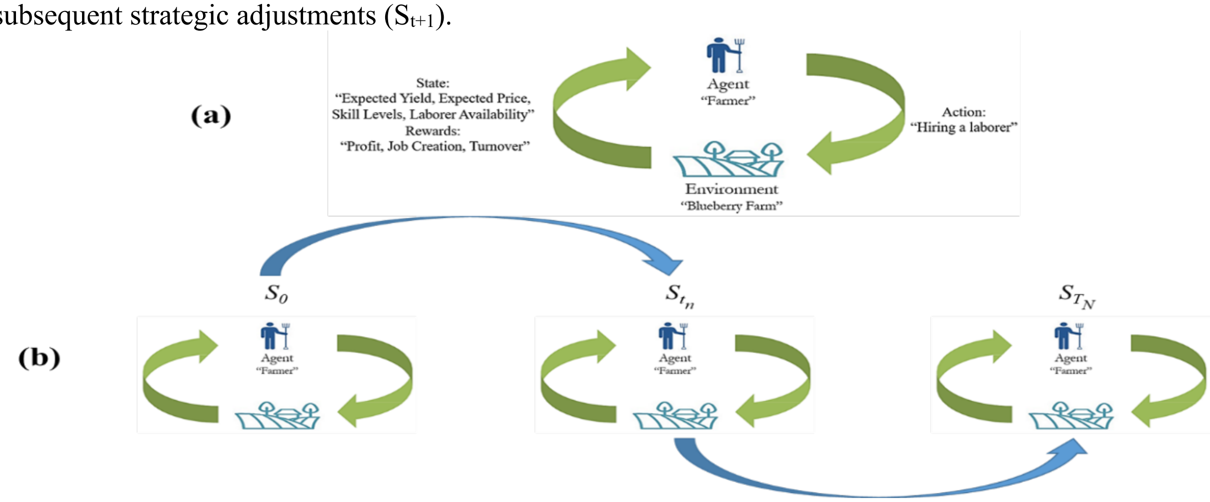

Figure 1: Farmer's action and learning process; (a) one state, (b) multiple states in one year and multiple years.

Amplifying the complexity of this decision-making problem is the categorization of labour into three skill levels, beginner (k=1), intermediate (k=2), and advanced (k=3), each associated with distinct productivity and quality impacts on the harvesting process. The selection of labour, hence, not only influences the immediate yield quantity and quality but also reflects the farm's commitment to sustainable practices and social responsibility. This intricate decision network, underscored by fluctuating market conditions, environmental variability, and socio-economic considerations, necessitates an analytical framework capable of processing high-dimensional data and adapting to dynamic changes.

The planning time horizon *N* considered in study is 20 years during which a farmer makes strategic human resource management decisions on a yearly basis for blueberry farming. The 20-year period differentiates short-term decisions, which prioritize immediate gains, from long-term planning that focuses on cumulative impacts, sustainability, and strategic investments for lasting success and maximized returns over two decades. Given the dynamic, complex, and multi-objective nature of sustainable HRM in agriculture, traditional *Deep Learning* (DL) and *Modeling and simulation* (M&S) approaches exhibit certain limitations. While proficient in identifying patterns within large datasets, DL models need more temporal depth and decision-sequential understanding, which are required for the iterative decision-making processes characteristic of HRM strategies. Conversely, M&S approaches, although valuable for hypothesis testing and scenario exploration, may not adequately capture the real-time dynamism and unpredictability inherent in agricultural environments. DRL, based on iterative learning and decision-making from environmental interactions, emerges as a methodologically superior choice for this study.

DRL's capacity to navigate complex decision-making processes, adaptability, and continuous learning capabilities align perfectly with sustainable HRM requirements in agriculture [18, 19]. Moreover, DRL's proficiency in handling multi-objective optimization and high-dimensional state spaces [20, 21] supports its application in balancing sustainability's economic, social, and environmental dimensions in labour hiring practices. DRL's practical relevance and adaptability to agricultural settings further justify its selection, promising actionable insights, and strategies for sustainable HRM. In addition, DRL integrates deep learning with reinforcement learning, enabling the handling of large state and action spaces that would otherwise be intractable with conventional RL techniques. This capability is crucial in sustainable HRM problems, where the decision space encompasses a wide range of variables including crop yields, labor skill levels, and market prices. For example, in this study we consider 2^{15} possible actions at each state, and ${2^{15}}^{20}$ potential policy paths. The deep learning component of DRL allows for the extraction of high-level feature representations from raw data, making it possible to discern intricate patterns and relationships that inform more strategic and effective decision-making.

State representation: The state space consists of all possible states, i.e. $\mathcal{D}$, during harvesting, characterized by four inputs at time step t. In this study, a state at time step *t* is represented by a vector $\vec{S}^t = [\tilde{y}^t, \tilde{p}^t, w_i^t, a_i^t] \in \mathcal{D}$. In this vector, $\tilde{y}^t$ represents the expected yield at time *t*, $\tilde{p}^t$ represents the expected price at time t, w_i^t refers to the skill levels of laborers at time t, which are randomly defined at the beginning of harvesting and may improve if a worker is selected for training, and a_i^t is a binary variable that determines a laborer's availability at time *t*.

Action space: The action space, i.e. $\mathcal{A}$, comprises all potential decision-actions a farmer might take at each time step t. The decision at time step *t* is defined as a vector $\vec{X}^t = [x_i^t] \in \mathcal{A}$, where x_i^t is a binary variable and becomes 1 if worker *i* is hired and otherwise 0.

Learning and forgetting functions: Workers are categorized into beginner, intermediate, and advanced skill levels, with wages and efficiency varying accordingly. Higher-skilled workers earn more and deliver better quality work. Skill improvement over time is possible through learning, while a fairness mechanism, the forgetting function, may affect hiring choices [22, 23]. The functions for learning and forgetting, represented by (1), are applied to each worker yearly, with L_i^t and F_i^t respectively.

$$\begin{cases} L_i^t = \beta - \beta e^{-\lambda_i^t . t} & if x_i^t = 1, \\ F_i^t = e^{-\gamma_i^t . t} & otherwise. \end{cases} \tag{1}$$

The rates of learning and forgetting for worker *i* at time *t* are defined as λ_i^t and γ_i^t respectively. Additionally, β represents the maximum potential improvement in a worker's skill, calculated by subtracting the previous skill level from the highest attainable skill level, which is 1. Accordingly, the worker's current skill, i.e. $\theta(x_i^t)$, gets updated based on potential skill, learning from work experience (i.e. $\theta(x_i^t) = \theta(x_i^{t-1}) + L_i^{t-1}$) and the impact of time away from work for each period (i.e. $\theta(x_i^t) = \theta(x_i^{t-1}) - F_i^{t-1}$).

3.1 Economic Model

To calculate the economic reward function, the farmer evaluates revenue, labor costs, and missed opportunities to maximize profits from selling harvested blueberries. This consideration not only aims at optimizing worker wages and mitigating potential yield losses but also supports environmental objectives. The profit function, as detailed in (2), Emphasizes the critical need to minimize product waste through the reduction of missed opportunities, thereby integrating environmental considerations into this function.

$$\Pi^t = (\tilde{p}^t . h^t) - (\vec{X}^t . \vec{W}^t) - (\tilde{p}^t . (\tilde{y}^t - h^t)), \tag{2}$$
$$\text{s.t. } h^t \leq \tilde{y}^t \quad \forall t,$$

where, $\tilde{p}^t$ represents the price at time *t*, h^t denotes the quantity of harvested blueberries at time *t*, where $h^t = (\sum_i x_i^t . \theta(x_i^t)) \times$ (Highest Harvest), where the highest harvest is calculated based on the maximum harvested value by advanced labor with $\theta(x_i^t) = 1$. Also, $\tilde{y}^t$ determines the total possible harvested blueberry at time *t*, and $\vec{X}^t$ and $\vec{W}^t$ are vectors of worker x_i^t relative to their corresponding wage w_i^t. Besides, (3) determines the normalized economic reward function in farming, calculated as the ratio of profit to income. In the economic function, the first term calculates the total revenue earned from selling harvested blueberries at a specific price. The second term measures the wages paid to workers based on their skill levels. Using the third term, lost opportunities are determined by calculating the difference between the actual yield and the harvested blueberries; this value shows how many crops could have been harvested during the period.

$$ER^t = \frac{\Pi^t}{\tilde{p}^t . \tilde{y}^t}. \tag{3}$$

3.2 Social Model

Furthermore, the farmer places a high priority on social responsibilities, focusing on boosting job creation and reducing turnover. This approach includes retaining existing teams for subsequent periods or hiring new labor only as needed. (4) and (5) quantify social activity values, assessing job creation and turnover, respectively. (4) yields a normalized value for job creation, considering the total pool of available workers, while (5) calculates the normalized value of turnover by comparing it with the figures from the previous time step.

$$SR1^t = \frac{Card\left(\overrightarrow{X}^t \middle| x_i^t = 1\right)}{\sum_i a_i^t}; \tag{4}$$

$$SR2^t = \frac{\sum_i a_i^t - \left|\overrightarrow{X}^{t+1} - \overrightarrow{X}^t\right|}{\sum_i a_i^t}. \tag{5}$$

Therefore, (6) defines the social reward function, indicating that the ideal values for (4) and (5) are 1, signifying maximum job creation and minimum turnover. Consequently, a multiplication operator is employed to compute the final value. To put it simply, we want to create more jobs and keep turnover low at the same time. Indeed, outcomes that either yield low job creation coupled with low turnover, or high job creation with high turnover, are deemed unsatisfactory. Therefore, employing multiplication as a strategy effectively addresses this challenge, ensuring both objectives are met concurrently.

$$SR^t = SR1^t \times SR2^t. \tag{6}$$

3.3 Total Utility

We utilize *Multi-Attribute Utility Theory* (MAUT), as outlined by [24] and represented in (7), for the aggregation of utilities. The elasticity coefficients are derived using elicitation techniques, which involve the comparative evaluation of payoff lotteries, as described by [17]. This method provides the versatility to examine various scenarios, spanning from purely economic to entirely social perspectives, by adjusting the weights of each objective, with the condition that $\mu_E + \mu_S = 1$, where μ_E and μ_s show the weight of economic and social objectives, respectively. This allows for a comprehensive comparison of outcomes and the extraction of pertinent managerial insights.

$$U(R^t) = \mu_E . U(ER^t) + \mu_S . U(SR^t). \tag{7}$$

To ensure a comparative analysis between short-term and long-term strategies, we will employ the Maskable PPO algorithm, adjusting the discount factor accordingly. This algorithm enhances performance by masking invalid actions in each state, preventing their selection by the agent. For short-term focus, we will set the discount factor λ to approximately 0, directing the agent to prioritize immediate rewards. Conversely, for the long-term approach, λ will be set to approximately 1, encouraging the agent to weigh long-term consequences more greatly. We assume the agent's state encompasses expected yield, expected price, individual worker skill levels, and laborer availability. The action involves selecting whom to hire, i.e. 0 (not hire) or 1 (hire), from a pool of workers, naturally excluding those who are unavailable. The reward is derived from the utility function $U(R^t)$, representing the utility of rewards at time t. We utilize a specialized DRL algorithm for blueberry farming HRM, detailed in Algorithm 1 (see Appendix A), considering yield, price, farmer preferences, and initializing policy and value functions. This algorithm simulates farm operations, updating worker attributes and making hiring decisions based on current conditions. It evaluates economic and social outcomes to balance profit and social responsibility. Through iterations, it learns an optimized HR policy considering economic, social, and time goals. We assess strategy effectiveness by comparing trained agents across time horizons and farmer personas.

4 RESULTS

In this study, we leveraged the StableBaselines3-Contrib package, a sophisticated tool in Python designed for training *Reinforcement Learning* (RL) models. This choice was driven by its comprehensive suite of algorithms and compatibility with complex simulation environments. Our computational setup comprised a high-performance computing system equipped with a 13th Generation Intel® Core™ i9-13900HX processor, boasting a base operating frequency of 2.20 GHz and augmented by 32.0 GB of RAM. This powerful hardware configuration facilitated the intensive computational demands of our RL models. The training process for each environment and agent was conducted extensively, lasting approximately one day per session, utilizing 1 million timesteps. This rigorous training regimen was necessary to ensure that our models could accurately learn and adapt to the complex dynamics of agricultural decision-making in the context provided.

Our case study imagines a farmer managing a 30-acre blueberry farm in British Columbia over 20 years. With 15 workers annually, each with varying availability and skills, the farmer aims to optimize productivity and profitability. Data on production, farm value, and worker wages in British Columbia (statcan.gc.ca) inform decision-making (see Table A). Through computational experiments, we assess the effectiveness of RL techniques in agricultural management for more sustainable practices. We created farmer personas based on economic and social priorities and short or long-term planning. *Short-Term* (ST) personas prioritize immediate financial returns, often at the cost of long-term sustainability and social welfare. In contrast, *Long-Term* (LT) personas focus on adopting practices that ensure the farm's long-term viability, emphasizing sustainability and broader social benefits. With a preference for economic outcomes, *Economically Centered* (EC) personas strive for a 0.75 to 0.25 economic-to-social objective ratio, aiming to maximize financial returns potentially at the expense of social endeavours. *Socially Oriented* (SO) personas, however, prioritize social goals over pure economic gain, operating with a 0.75 to 0.25 social-to-economic ratio. Our experiments, outlined in Table 1, highlight the advantages of long-term farming strategies over short-term ones. Long-term planning leads to higher profits, better turnover rates, improved workforce management, and stable operations. These benefits optimize resource use and maintain efficiency over time, fostering sustainable practices and economies of scale for better returns on investment. A stable workforce is crucial for long-term success, promoting skill development, loyalty, and a positive work environment. By prioritizing workforce stability, farms reduce recruitment costs and cultivate a motivated team, boosting operational efficiency. Investing in human capital through continuous learning sustains long-term strategies, enhancing competitiveness and adaptability to market changes. Long-term planning creates sustainable job opportunities, reducing the need for frequent hiring and training. Comparing short-term and long-term strategies for EC and SO personas, long-term planning significantly reduces turnover by over 80% and boosts profits by at least 90%. While short-term planning may create more jobs initially, it falls short of meeting socially responsible objectives, highlighting the importance of long-term strategic planning in balancing financial success and social responsibility in agriculture. SO personas experience slightly higher turnover rates due to their strategy of hiring all available labor, increasing the risk of unavailability in subsequent periods.

Table 1: Pareto results of economic and social objectives from 1,000 simulations.

Planning Horizon	**Persona**	**Job Creation**		**Turnover**		**Profit**	
		Avg.	**Std.**	**Avg.**	**Std.**	**Avg.**	**Std.**
Short-term	SO	161	12.56	126	10.45	198,705.19	173,263.30
	EC	127	9.18	102	9.47	369,233.95	74,055.36
Long-term	SO	156	8.78	22	4.7	489,444.49	17,4124.70
	EC	119	8.37	17	4.35	702,756.09	71,734.26

5 CONCLUSIONS

This study highlighted the effectiveness of DRL in assessing the relative benefits of long-term versus short-term decision-making strategies in the agri-food supply chain, focusing particularly on blueberry harvesting in British Columbia, Canada. Through detailed comparative analysis, the superiority of long-term strategies became apparent, showcasing their ability to not only increase profitability for farmers but also foster better social outcomes. By dividing sustainable decisions into economic and social dimensions, this research emphasized the necessity of aligning profit with social benefits, such as job creation and turnover reduction, within the agricultural sector. DRL's application enabled an in-depth examination of sustainable decision-making, with implications reaching beyond to the wider agri-food supply chain. In essence, the findings strongly support the preference for long-term strategies over short-term ones, underlining their advantage in terms of both profitability and social impact. This confirmation of DRL's utility in strategic agricultural decision-making highlights the critical need for adopting long-term approaches to ensure sustainable agricultural practices and societal welfare.

A APPENDICES

The pseudocode for the algorithm is as follows.

ALGORITHM 1: DEEP REINFORCEMENT LEARNING OF HUMAN RESOURCE MANAGEMENT.

```
Input: Blueberry yield, Blueberry price, Farmer's economically driven weight, Farmer's social responsibility weight, Farmer's decision-making time horizon
Output: Learned policy π
Initialize policy parameters θ
Initialize value function parameters ϕ
For iteration = 1, 2, … Number of Training; do
    Initialize Worker's availability, Worker's skill, Worker's wage
    While not done
        Observe the current state
        Select an action on the current policy which is a vector of binary variables showing whom to hire for the current year (1 is hire)
        Execute the action in the environment
        Calculate the economic and social rewards
        Calculate the best (Max) and worst (Min) possible economic and social rewards
        Normalize economic and social rewards
        Calculate utility function
        Update each worker's skill based on the learning and forgetting function
        Update each worker's wage based on their skill
        Observe the next state and whether the episode is done
        Store the transition
    end
    Update the policy and value networks
end
```

The inputs used in the DRL training are presented in Table A as follows.

Table A: Dataset.

Parameter	Value
Total timestep and Farm's area	20 years & 30 acres
Highest harvest	1,5246 kg/harvesting season
Number of workers in the pool	15 workers
[Advance; Intermediate; Beginner] wages	[22,500; 13,500; 9,000] \$/year
Hiring cost	5,000 \$/worker
Learning and Forgetting rates	$N(0.1, 0.05)$; $N(0.05, 0.025)$
Probability of workers' availability	0.6
Skill range of advance workers	(0.85,1]
Skill range of Intermediate workers	(0.65, 0.85]
Skill range of Beginner workers	[0.4, 0.65]

REFERENCES

[1] Food and Agriculture Organization of the United Nations (FAO), "Greenhouse gas emissions from the dairy sector: a life cycle assessment," 2010. [Online]. Available: https://www.fao.org/4/k7930e/k7930e00.pdf

[2] J. A. Foley *et al.*, "Solutions for a cultivated planet," *Nature,* vol. 478, no. 7369, pp. 337-342, 2011.

[3] W. Schlenker and M. J. Roberts, "Nonlinear temperature effects indicate severe damages to US crop yields under climate change," *Proceedings of the National Academy of sciences,* vol. 106, no. 37, pp. 15594-15598, 2009.

[4] D. A. Ankrah, J. Mensah, J. N. Anaglo, and S. D. Boateng, "Climate variability indicators-scientific data versus farmers perception; evidence from southern Ghana," *Cogent Food & Agriculture,* vol. 9, no. 1, p. 2148323, 2023.

[5] C. B. Field, *Managing the risks of extreme events and disasters to advance climate change adaptation: special report of the intergovernmental panel on climate change*. Cambridge University Press, 2012.

[6] P. K. Thornton, P. J. Ericksen, M. Herrero, and A. J. Challinor, "Climate variability and vulnerability to climate change: a review," *Global change biology,* vol. 20, no. 11, pp. 3313-3328, 2014.

[7] S. Seuring and M. Müller, "From a literature review to a conceptual framework for sustainable supply chain management," *Journal of cleaner production,* vol. 16, no. 15, pp. 1699-1710, 2008.

[8] Y. Shan *et al.*, "Optimizing the maize irrigation strategy and yield prediction under future climate scenarios in the Yellow River Delta," *Agronomy,* vol. 13, no. 4, p. 960, 2023.

[9] D. Elavarasan and P. D. Vincent, "Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications," *IEEE access,* vol. 8, pp. 86886-86901, 2020.

[10] G. E. Ogutu, W. H. Franssen, I. Supit, P. Omondi, and R. W. Hutjes, "Probabilistic maize yield prediction over East Africa using dynamic ensemble seasonal climate forecasts," *Agricultural and forest meteorology,* vol. 250, pp. 243-261, 2018.

[11] U. Lehtinen, "Sustainable supply chain management in agri-food chains: a competitive factor for food exporters," *Sustainability Challenges in the agrofood sector,* pp. 150-174, 2017.

[12] B. S. Nichols, H. Stolze, and J. F. Kirchoff, "Spillover effects of supply chain news on consumers' perceptions of product quality: An examination within the triple bottom line," *Journal of Operations Management,* vol. 65, no. 6, pp. 536-559, 2019.

[13] S. Matos and J. Hall, "Integrating sustainable development in the supply chain: The case of life cycle assessment in oil and gas and agricultural biotechnology," *Journal of operations management,* vol. 25, no. 6, pp. 1083-1102, 2007.

[14] R. Narasimhan, A. Nair, D. A. Griffith, J. S. Arlbjørn, and E. Bendoly, "Lock-in situations in supply chains: A social exchange theoretic study of sourcing arrangements in buyer–supplier relationships," *Journal of Operations Management,* vol. 27, no. 5, pp. 374-389, 2009.

[15] S. Ambulkar, J. Blackhurst, and S. Grawe, "Firm's resilience to supply chain disruptions: Scale development and empirical examination," *Journal of operations management,* vol. 33, pp. 111-122, 2015.

[16] B. Jiang, R. C. Baker, and G. V. Frazier, "An analysis of job dissatisfaction and turnover to reduce global supply chain risk: Evidence from China," *Journal of operations management,* vol. 27, no. 2, pp. 169-184, 2009.

[17] T. Dvir and Y. Berson, "Leadership, motivation, and strategic human resource management," *Handbook of industrial engineering: technology and operations management,* pp. 837-867, 2001.

[18] V. Mnih *et al.*, "Human-level control through deep reinforcement learning," *nature,* vol. 518, no. 7540, pp. 529-533, 2015.

[19] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, and M. Riedmiller, "Deterministic policy gradient algorithms," in *International conference on machine learning*, 2014: Pmlr, pp. 387-395.

[20] K. Van Moffaert and A. Nowé, "Multi-objective reinforcement learning using sets of pareto dominating policies," *The Journal of Machine Learning Research,* vol. 15, no. 1, pp. 3483-3512, 2014.

[21] T. P. Lillicrap *et al.*, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971,* 2015.

[22] M. J. Anzanello and F. S. Fogliatto, "Learning curve models and applications: Literature review and research directions," *International Journal of Industrial Ergonomics,* vol. 41, no. 5, pp. 573-583, 2011.

[23] P. Woźniak, E. Gorzelańczyk, and J. Murakowski, "Two components of long-term memory," *Acta neurobiologiae experimentalis,* vol. 55, no. 4, pp. 301-305, 1995.

[24] R. L. Keeney and H. Raiffa, *Decisions with multiple objectives: preferences and value trade-offs*. Cambridge university press, 1993.

AUTHORS BIOGRAPHIES

VAHID GHOLAMZADEH is a Ph.D. student in an interdisciplinary program between Computer Science and the Gustavson School of Business, University of Victoria, Canada. His email address is vahidgh@uvic.ca.

ADEL GUITOUNI is an associate professor of management sciences, decision support systems and international business at the Gustavson School of Business, University of Victoria, Canada. His email address is adelg@uvic.ca.

BEHROOZ KHORSHIDVAND is a post-doctoral researcher at the Gustavson School of Business, University of Victoria, Canada. His email address is behroozkhorshidvand@uvic.ca.