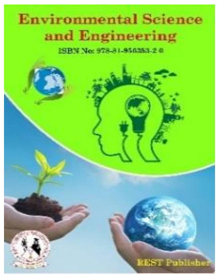
Environmental Science and Engineering
Vol: 2(3), 2023
REST Publisher; ISBN: 978-81-956353-2-0

Website: http://restpublisher.com/book- series/environmental-science-and-engineering/

DOI: https://doi.org/10.46632/ese/2/3/1



Promoting Green Supply Chain Management With Optimal Selection Of Packaging Materials Using Integrated Fuzzy Mcdm And Rl Model

[1]P. Pandiammal, [2]N. Ramila Gandhi, *[3]Nivetha Martin
[1]*GTN Arts College Autonomous, Dindigul, TamilNadu, India.*
[2]*PSNA College of Engineering and Technology Autonomou, Dindigul, TamilNadu, India.*
[3]*Arul Anandar College Autonomous, Dindigul, TamilNadu, India.*
*Corresponding Author Email: nivetha.martin710@gmail.com

Abstract: *Green supply chain management is highly significant to maintain environmental sustainability. The agglomeration of green components enhances and supports the business activities to practice green supply chain more effectively. Utilizing sustainable packaging materials in logistics is a step towards promoting business eco sustainability. This research work attempts to develop a hybrid decision making model by integrating techniques of fuzzy multi criteria decision making (MCDM) and Reinforcement Learning (RL). This research work proposes a decision-making method of IDOCRIW (Integrated Determination of Objective Criteria Weights) under fuzzy environment with linguistic representations to determine the criterion weights of material selection and applies the RL method of Q learning in ranking the packaging materials for promoting green sustainability. The proposed fuzzy based MCDM method resolves the problems of conflict of uncertainty. The ranking results obtained using this method are compared with the non-integrated MCDM method. The proposed combined model shall be discussed under various other extended fuzzy representations. The decision-making problem on optimal selection of packaging materials addressed in this research work benefits the business decision makers to make right choices. This hybrid model will certainly make the logistic environment more robust and also it will upscale the smart framework of supply chain management.*
Keywords: *Green supply chain management, fuzzy, IDOCRIW, Reinforcement learning, Q learning*

1. INTRODUCTION

Green Supply chain management (GSCM) is a collection of environmental oriented activities that intends to alleviate the environmental impacts during the integrated processs of supply chain. The practice of GSCM benefits in cost saving, regulatory compliance and competitive advantage. Some of the key practices of GSCM are green procurement, energy efficiency, waste reduction, reverse logistics and carbon footprint reduction. One of the most significant aspects that contributes to the effective functioning of GSCM is the materials used for packaging. The packaging materials vary based on its nature, characteristics and functions. Why is GSCM highly concerned about packaging materials? The reason is very direct, it is the reflection of environmental commitment and also the alignment with the stakeholders to conserve the environment from the mishaps of wastes. In a GSCM framework, the packaging materials are chosen not only based on the Eco-friendliness but also on the other attributes of affordability, functionality, productivity, robustness, adaptability, versatility, durability. The material selection becomes optimal by using a multi-criteria decision-making process. A MCDM method in general falls into either of the two categories of criterion weight computation and alternatives ranking. The MCDM methods are more scientifically designed to arrive at the ideal solutions by combining with the existing different methods of decision making. In recent times the different methods of MCDM are combined with ML based algorithms to develop hybrid and integrated models. In this research work a more integrated decision-making method is developed by combining the method of fuzzy IDOCRIW with the RL algorithm of Q learning. As the decision-making environment involves uncertainty, the need of formulating fuzzy centric methods is inevitable and this research work is a step towards it. The integrated models of the above mentioned will be of much significance in dealing with the complicated problems of making decisions. The remaining content of the paper is structured as follows; section 2 presents the detailed contributions of research works in the proposed area of decision making. Section 3 sketches the steps involved in fuzzy IDOCRIW and RL. Section 4 applies to the proposeddecision making problem and discusses the results. Section 5 concludes the work.

2. LITERATURE REVIEW

This section presents the review of works associated with the decision making on packaging materials in the context of supply chain management. Manupathi [1] applied a fuzzy Analytic Hierarchy Process in determining the drivers of green manufacturing with respect to packaging materials. Tavana et al [2] employed type-2 fuzzy based best-worst method and combined compromise solution for evaluating eco-friendly packaging alternatives. Gurrala et al [3] applied an integrated CRITIC and TOPSIS method to select edible packaging materials. Bingol [4] used Fuzzy AHP-Entropy and proximity index value method in selecting semiconductor packaging materials. Ghosh et al [5] evaluated the performance of GSCM using fuzzy representations. Ling et al [6] developed a framework of GSCM in a hybrid sense. Muenjitnoy and Sonthipermpoon [7] applied AHP in selecting food packaging materials. Mahajan et al [8] applied integrated MCDM in selecting sustainable materials. Xie et al [9] employed a three stage fuzzy network and DEA method. Different combinations of MCDM methods are used in selecting optimal packaging materials based on different criterion weights. In general the criterion weights play a significant role in ranking the alternatives. The most commonly applied methods to find the criterion weights are method of Entropy, AHP. However, an improved version of the method of finding the criterion weight was developed by Zavadskas and Podvezko [10] in 2016. The method of Entropy and Criterion Impact LOSs is combined together to frame a blended model by name Integrated Determination of Objective CRIteria Weights. This newly framed model is applied by binding with other MCDM methods of ranking. Alao et al [11] integrated with TOPSIS, Luo et al [12] with COCOSO, Zarch et al [13] with WASPAS, Pala [14] with MARCOS. Also the methods of MCDM are integrated with machine learning algorithms to ease the decision making process. The machine learning algorithms are classified into supervised, unsupervised, reinforcement and deep learning. Reinforcement Learning is a machine learning paradigm where an agent learns to make sequential decisions by interacting with an environment. Using the RL algorithm, Guan et al [15] modelled a behaviour based decision making model, Maqsood et al [16] developed an autonomous framework, Singi et al [17] formulated robotic agents based decision making models. RL algorithms are widely applied in decision making but not much integrated with MCDM methods. From the above described literature the following shortcomings are identified. The RL algorithms are not widely blended with MCDM methods RL algorithms, especially Q learning techniques, are not applied in decision making on packaging materials in the context of GSCM. In this research work a more integrated method of combining fuzzy based IDOCRIW and Q learning is developed.

3. METHODOLOGY OF INTEGRATED FUZZY IDOCRIW- Q LEARNING

The steps involve in the proposed integrated method is presented as follows in Fig.1 . This method comprises two phases, the first phase is finding of criterion weights and the second phase is ranking of alternatives

Phase I: Fuzzy **IDOCRIW**

Step 1: Formulation of Decision Making matrix with linguistic values associating the alternatives and the criteria. $D = \begin{bmatrix} L_{11} & \cdots & L_{in} \\ \vdots & \ddots & \vdots \\ L_{n1} & \cdots & L_{nn} \end{bmatrix}$

Step 2: Normalization of the Matrix:

$\overline{L_{ij}} = \frac{L_{ij}}{\sum_{i=1}^{n} L_{ij}}$, j = 1,..,n

Step 3: Finding the Degree of Entropy:

$$E_j = \frac{1}{\ln n} \sum_{i=1}^{n} L_{ij} * \ln L_{ij} \quad 0 \le E_j \le 1$$

Step 4 : Finding the Entropy weight

$d_j = 1 - E_j,\ w_j = \frac{d_j}{\sum_{j=1}^{n} d_j}$

Step 5 : Formulating a square matrix

$$\widehat{L_{ij}} = \frac{\min L_{ij}}{L_{ij}}$$

Step 6: Finding Relative Impact loss

$p_{ij} = \frac{a_{jj} - a_{ij}}{a_{jj}}$, a_j = max L_{ij}

Step 7: Formulating weight system matrix $\begin{bmatrix} -\sum_{i=1}^{n} P_{i1} & \cdots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{n1} & \cdots & -\sum_{i=1}^{n} P_{in} \end{bmatrix}$

Step 8: Finding the final weights

$$g_j = \frac{q_{j*w_j}}{\sum_{j=1}^{n} q_{j*w_j}}$$

Phase II: Q Learning

Q-learning is a popular reinforcement learning algorithm used for making decisions in environments with discrete states and actions. It's based on the idea of learning an action-value function (Q-function) that estimates the expected cumulative reward when taking a particular action in a specific state. Here are the steps involved in Q-learning:

Step 1: Initialization:

- Initialize the Q-values for all state-action pairs arbitrarily, or to some initial values.
- Set hyperparameters like the learning rate (α) and the discount factor (γ).

Step 2: Exploration vs. Exploitation:

- Choose an exploration strategy, such as ε-greedy or softmax, that determines whether the agent explores new actions or exploits the current best action based on Q-values.

Step 3: Episode Loop:

- Start an episode by initializing the environment and setting the initial state.
- Enter a time-step loop within the episode.

Step 4 : Action Selection:

- Use the exploration strategy to choose an action based on the current state and Q-values.

Step 5 : Interaction with Environment:

- Take the chosen action and observe the next state and the associated reward from the environment.

Step 6: Q-Value Update:

- Update the Q-value of the current state-action pair using the Q-learning update rule:

  Q(s, a) = Q(s, a) + α * [reward + γ * max(Q(next_state, all_actions)) - Q(s, a)]

  - **α** is the learning rate.
  - γ is the discount factor, representing how much the agent values future rewards.
  - **max(Q(next_state, all_actions))** estimates the best expected reward in the next state.

Step 7: State Transition:

- Move to the next state based on the environment's response to the chosen action.

Step 8: Termination:

- Repeat the time-step loop until the episode terminates, either due to a terminal state or a predefined number of steps.

Step 9: Episodic or Continuous Learning:

- If the problem is episodic (e.g., games), repeat the episode loop for a certain number of episodes.
- For continuous tasks, the agent can interact with the environment indefinitely.

Step 10: Convergence:

- Over time, as the agent interacts with the environment, the Q-values converge to the optimal action-values.

Step 11: Policy Extraction:

- The learned Q-values can be used to extract an optimal policy by selecting the action with the highest Q-value for each state.

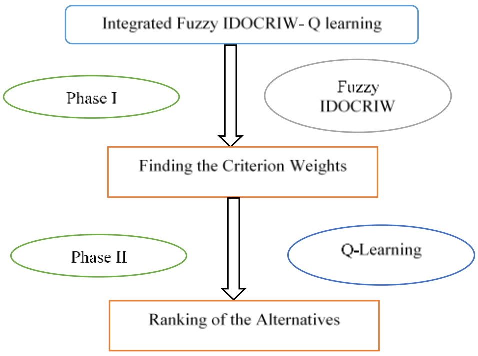


FIGURE 1. Overall Framework of the Proposed Method

4. APPLICATION OF FUZZY IDOCRIW- Q LEARNING ALGORITHM IN OPTIMAL SELECTION

This section applies the proposed integrated method in ranking the alternatives of packaging materials. A hypothetical example is developed with 5 alternatives and 5 criteria. The initial decision-making matrix is presented as follows in Table 1.

TABLE 1. Initial Decision-Making Matrix

Packaging Materials/ Criteria	**C1**	**C2**	**C3**	**C4**	**C5**
F1	M	L	M	H	H
F2	M	M	L	L	VL
F3	H	H	VH	VH	VL
F4	L	M	M	H	H
F5	VL	VL	M	M	H

The criteria considered for making decisions is presented in Table 2

TABLE 2. Criteria Description

C1	Adaptability
C2	Versatility
C3	Recyclable
C4	Durability
C5	Affordability

The linguistic variable are quantified using triangular fuzzy numbers is presented in Table 3

TABLE 3. Modified Matrix

Packaging Materials/ Criteria	**C1**	**C2**	**C3**	**C4**	**C5**
F1	0.45	0.23	0.45	0.84	0.84
F2	0.45	0.45	0.23	0.23	0.07
F3	0.84	0.84	0.95	0.95	0.07
F4	0.23	0.45	0.45	0.84	0.84
F5	0.07	0.07	0.45	0.45	0.84

By using the above integrated method the criterion weights are obtained using Fuzzy IDOCRIW and it is presented in Table 4

TABLE 4. Criterion Weights

C1	0.268674
C2	0.340886
C3	0.170225
C4	0.108349
C5	0.111866

Using Q learning in R programming software the ranking of the alternatives is obtained

```
# Q-learning implementation to rank 5 suppliers based on 5 criteria
# Define the number of suppliers and criteria
num_suppliers <- 5
num_criteria <- 5
# Initialize Q-table with zeros
num_states <- num_suppliers ^ num_criteria
```

```
num_actions <- num_suppliers
Q <- matrix(0, nrow = num_states, ncol = num_actions)
# Define rewards (lower is better for simplicity, adjust as needed)
rewards <- matrix(c(
  100, 80, 60, 40, 20,
  80, 60, 40, 20, 100,
  60, 40, 20, 100, 80,
  40, 20, 100, 80, 60,
  20, 100, 80, 60, 40
), ncol = num_suppliers, byrow = TRUE)
gamma <- 0.8  # Discount factor
learning_rate <- 0.1
num_episodes <- 1000
max_steps <- 100
# Q-learning algorithm
for (episode in 1:num_episodes) {
  state <- sample(1:num_states, 1)  # Start in a random state
    for (step in 1:max_steps) {
    action <- sample(1:num_actions, 1)  # Choose a random action (supplier)
    next_state <- action  # In this example, actions transition to next states
    reward <- rewards[state, action]  # Get the reward for the next state
        # Q-learning update rule
    Q[state, action] <- Q[state, action] + learning_rate * (reward + gamma * max(Q[next_state, ]) - Q[state, action])

    state <- next_state  # Move to the next state
        if (step == max_steps) {
      break  # Reached the maximum number of steps
    }
  }
}
# Rank suppliers based on the learned Q-values
supplier_rank <- order(-Q[1, ])  # Starting state is 1
print("Ranked Suppliers:")
print(supplier_rank)
```

By using the above code the ranking results of the alternatives are obtained as in Table 5

TABLE 5. Ranking results using Q-learning

F1	F2	F3	F4	F5
3	1	2	4	5

The ranking results obtained using other integrated MCDM methods is presented in Table 6

TABLE 6. Comparative Analysis of the ranking results

Methods / Ranking	F1	F2	F3	F4	F5
Fuzzy IDOCRIW& Q Learning	3	1	2	4	5
Fuzzy IDOCRIW &TOPSIS	3	1	2	4	5
Fuzzy IDOCRIW &MARCOS	3	1	2	4	5

The ranking results justifies the efficacy of the results obtained using the proposed method. The results obtained using other methods also coincides with the obtained result.

5. CONCLUSION

This research work proposes a hybrid decision making model using the method of fuzzy IDOCRIW & Q Learning in selecting packaging materials in Green Supply chain management. The proposed method is so robust by nature as it yields optimal ranking results. The numerical example based on hypothetical data substantiates the efficiency of the hybrid method. The proposed method has high business utility value especially in making the green supply chain management more effective.

REFERENCES

[1]. Manupati, V. K. (2022). Prioritization of Drivers for Lean and Green Manufacturing Using Fuzzy Analytic Hierarchy Process (FAHP) Approach. Advances in Manufacturing Processes, Intelligent Methods and Systems in Production Engineering: Progress in Application of Intelligent Methods and Systems in Production Engineering, 335, 117.

[2]. Tavana, M., Shaabani, A., Di Caprio, D., &Bonyani, A. (2022). A novel Interval Type-2 Fuzzy best-worst method and combined compromise solution for evaluating eco-friendly packaging alternatives. Expert Systems with Applications, 200, 117188.

[3]. Gurrala, K. R., Helmy, M., &Ndiaye, M. (2022, March). Edible Packaging Selection Employing Hybrid CRITIC and TOPSIS Method. In 2022 International Conference on Decision Aid Sciences and Applications (DASA) (pp. 822-826). IEEE.

[4]. Bingol, S. (2022). Selection of Semiconductor Packaging Materials by Combined Fuzzy AHP-Entropy and Proximity Index Value Method. Mathematical Problems in Engineering, 2022.

[5]. Ghosh, S., Mandal, M. C., & Ray, A. (2023). A PDCA based approach to evaluate green supply chain management performance under fuzzy environment. International Journal of Management Science and Engineering Management, 18(1), 1-15.

[6]. Ling, L., Anping, R., & Di, X. (2023). Proposal of a hybrid decision-making framework for the prioritization of express packaging recycling patterns. Environment, Development and Sustainability, 25(3), 2610-2647.

[7]. Muenjitnoy, A., &Sonthipermpoon, K. (2023). The Food Packaging Design and Optimization of Small Business Entrepreneurs Based on Analytic Hierarchy Process. Scandinavian Journal of Information Systems, 35(1), 1000-1008.

[8]. Mahajan, A., Singh, I., & Arora, N. (2023). An integrated multi-criteria decision-making framework for the selection of sustainable biodegradable polymer for food packaging applications. Environment, Development and Sustainability, 1-22.

[9]. Xie, Y., Chen, N., Hua, G., & Xiao, Y. (2023). Methodology selecting packaging materials combined multi-sensory experience and fuzzy three-stage network DEA model. Packaging Technology and Science, 36(2), 125-134.

[10]. Podvezko, V., Zavadskas, E. K., &Podviezko, A. (2020). An Extension Of The New Objective Weight Assessment Methods Cilos And Idocriw To Fuzzy Mcdm. Economic Computation & Economic Cybernetics Studies & Research, 54(2).

[11]. Alao, M. A., Popoola, O. M., &Ayodele, T. R. (2021). Selection of waste-to-energy technology for distributed generation using IDOCRIW-Weighted TOPSIS method: A case study of the City of Johannesburg, South Africa. Renewable Energy, 178, 162-183.

[12]. Luo, Y., Zhang, X., Qin, Y., Yang, Z., & Liang, Y. (2021). Tourism attraction selection with sentiment analysis of online reviews based on probabilistic linguistic term sets and the IDOCRIW-COCOSO model. International Journal of Fuzzy Systems, 23, 295-308.

[13]. Eghbali-Zarch, M., Tavakkoli-Moghaddam, R., Dehghan-Sanej, K., &Kaboli, A. (2022). Prioritizing the effective strategies for construction and demolition waste management using fuzzy IDOCRIW and WASPAS methods. Engineering, Construction and Architectural Management, 29(3), 1109-1138.

[14]. Pala, O. (2021). IDOCRIW ve MARCOS TemelliBist Ulaştirmaİ şletmelerinin Finansal Performans Analizi. Kafkas Üniversites iİktis adiveİdari Bilimler Fakültesi Dergisi, 12(23), 263-294.

[15]. Guan, W., Zhao, M. Y., Zhang, C. B., & Xi, Z. Y. (2023). Generalized behavior decision-making model for ship collision avoidance via reinforcement learning method. Journal of Marine Science and Engineering, 11(2), 273.

[16]. Maqsood, M., Yasmin, S., Gillani, S., Aadil, F., Mehmood, I., Rho, S., & Yeo, S. S. (2023). An autonomous decision-making framework for gait recognition systems against adversarial attack using reinforcement learning. ISA transactions, 132, 80-93.

[17]. Singi, S., He, Z., Pan, A., Patel, S., Sigurdsson, G. A., Piramuthu, R., ...&Ciocarlie, M. (2023). Decision Making for Human-in-the-loop Robotic Agents via Uncertainty-Aware Reinforcement Learning. arXiv preprint arXiv:2303.06710.