*Human Factors in Robots, Drones and Unmanned Systems,* Vol. 179, 2025, 83–92
https://doi.org/10.54941/ahfe1006377

# Enhancing Trust in Human-AI Interaction Through Explainable Decision Support Systems for Mission Planning of UAS Swarms

**Batuhan Özcan, Max Friedrich, Jan-Paul Huttner, and Kevin Dwinger**

German Aerospace Center (DLR), Institute of Flight Guidance, Braunschweig, 38108, Germany

## ABSTRACT

Artificial Intelligence systems often operate as "black boxes", creating challenges for trust and collaboration in human-AI teaming environments. Particularly for Unmanned Aerial Systems swarm operations. This paper proposes an explainable decision support framework that integrates fuzzy-logic with reinforcement learning to enhance transparency, while maintaining adaptability. The framework processes uncertain inputs through linguistic variables and interpretable rules, generating natural language explanations alongside mission recommendations. Reinforcement learning optimizes system parameters within constraints, ensuring the decision-making process remains transparent while performance improves over time. This approach addresses key challenges in unmanned aircraft systems swarm coordination, particularly for dynamic task allocation when assets fail or environmental conditions change. By preserving explainability throughout the optimization process, the system enables operators to understand not only what decisions are made but why they are made, which is a crucial factor for establishing trust in human-autonomous teaming.

**Keywords:** Unmanned aircraft systems, Swarming, Mission management, Fuzzy logic, Artificial intelligence

## INTRODUCTION

The integration of Artificial Intelligence (AI) into decision-making systems has significantly advanced the capabilities of autonomous systems. However, the "black-box" nature of many AI models often raises concerns about transparency, interpretability and trustworthiness. These criteria's are especially important in safety-critical domains and the concerns are particularly pronounced in human-AI teaming, where trust plays a pivotal role ensuring effective collaboration and adoption of AI technologies. Human-AI teaming refers to the collaborative work between humans and AI systems in which both entities contribute complementary capabilities to achieve a shared goal. In this context humans typically provide contextual understanding and ethical judgement, while AI systems offer compu-tational power and data processing capabilities (Li et al., 2024; Thiebes et al., 2021; Seeber et al., 2018). Trust in AI systems is influenced by factors such as system reliability, transparency and the ability to provide meaningful explanations for decisions (Li et al., 2024; Mehrotra et al., 2023). Addressing these challenges is crucial to unlock the full potential of AI in the context of Unmanned Aerial Systems (UAS) swarms.

The simultaneous operation of multiple UAS, either independently or organized in coordinated swarms have emerged as an important technology in domains ranging from disaster response to surveillance and defence. The ResponDrone project demonstrates how drone-swarms can significantly enhance situational awareness and response capabilities in emergency scenarios (Polka et al., 2017; Erdelj et al., 2017). Their ability to operate collaboratively and adaptively offers unparalleled efficiency and scalability (Tahir et al., 2019). However, coordinating such a swarm involves complex decision-making under uncertainty. This requires a robust system, that can dynamically allocate tasks, manage resources and adapt to changing conditions (Yan et al., 2023; Lamont et al., 2007). It is important to note that current UAS-swarm technologies predominantly operate with a Human-In-The-Loop approach, rather than fully autonomy. The human operator remains essential for high-level decision-making, mission oversight and intervention in complex situations which automated systems cannot adequately handle alone. This human-machine interaction ensures appropriate ethical considerations and accountability while leveraging the computational advantages of automation (Endsley, 2016). Traditional AI approaches often struggle with the dual demands of high performance and explainability in these contexts (Wu and Xu, 2021). This limitation shows the need for an innovative framework, that balance adaptability with interpretability.

In this context, Fuzzy-Logic offers a compelling solution to this challenge by providing an inherently interpretable decision-making framework. Unlike classical AI methods that rely on non-transparent, statistical models, Fuzzy-Logic operates through intuitive linguistic rules and degrees of truth. Which makes it well-suited for modelling human-like reasoning in uncertain environments (Wu and Xu, 2021; Improta et al., 2019). For example, instead of rigid thresholds like “flight speed > 25 m/s”, Fuzzy-Logic employs terms such as “pretty fast” or “very fast”, which align more closely with human perception (Improta et al., 2019). This characteristic makes fuzzy logic particularly valuable in Human-Machine Interface (HMI) design, where promoting trust and collaboration is essential (Crandall and Cummings, 2007). Especially for mission management of multiple UAS or a UAS-swarm, studies have shown that trust is one of the most crucial human factors that need to be considered in the design of the HMIs (Friedrich, 2021). Current research in mission planning systems for UAVs confirm that establishing trust between operators and autonomous systems remains a significant challenge, that must be addressed through transparent interfaces and explainable decision processes (Huttner and Friedrich, 2023).

Building on these principles, this paper proposes a framework for an explainable decision support system for mission planning of UAS swarms. The approach proposes to integrate Fuzzy-Logic with Reinforcement Learning (RL) to optimize performance, while maintaining transparency.

Fuzzy-Logic serves as the foundation for modelling uncertainty and generating interpretable outputs, while RL enhances adaptability by fine-tuning the system parameters based on feedback from simulations (Melin and Castillo, 2013; Berenji and Khedkar, 1992; Kober et al., 2013). This hybrid approach ensures that the systems remains explainable even as it learns and evolves over time (Zander et al., 2023).

## BACKGROUND AND RELATED WORK

### Fuzzy-Logic in Decision Support Systems

Fuzzy-Logic has emerged as a powerful tool for decision-making in uncertain and complex environments. Unlike classical, binary logic, which operates in discrete true or false values, Fuzzy-Logic allows for reasoning with degrees of truth. This enables more of a human-like decision-making process. This flexibility makes it particularly well-suited for applications where data is incomplete, imprecise or ambiguous (Wu and Xu, 2021).

Fuzzy Decision Support Systems (FDSS) leverage linguistic variables, e.g. "high risk" or "moderate priority", and an intuitive rule-based framework to model human reasoning. These systems are inherently interpretable due to their reliance on transparent "if-then" rules, that align closely with human cognitive processes (Wu and Xu, 2021; Zander et al., 2023). For example, a Fuzzy-Rule might state: "If resource availability is low and task priority is high, then allocate additional resources".

However, challenges remain in subjective definition of membership functions and rules, as well as in integrating Fuzzy-Logic with data-driven approaches, like machine learning. Recent advancements have explored hybrid models combining Fuzzy-Logic with RL to address these limitations while maintaining system inter-pretability (Zander et al., 2023).

### UAS-Swarms and Mission Planning

UAS swarms represent a significant advancement in autonomous systems, offering scalability, redundancy and adaptability in dynamic environments (Tahir et al., 2019). These swarms are increasingly deployed in applications such as disaster response, surveillance and precision agriculture (Karampelia et al., 2023). However, coordinating multiple UAS introduces new challenges related to task allocation, resource management and real-time decision-making under uncertainty.

Dynamic task allocation is a critical component of swarm mission planning. Algorithms such as Ant Colony Optimization (ACO), Genetic Algorithms (GA) and Binary Wolf Pack Algorithms (BWPA) have been employed to optimize task distribution among UAVs. These strategies are based on factors like task priority, resource constraints and environmental conditions (Peng et al., 2021; 2022).

In addition to task allocation, UAS swarm coordination relies on robust communication protocols and formation control strategies to ensure efficient operation. Distributed control methods have proven superior to centralized approaches in enhancing swarm scalability and resilience to communication failures (Karampelia et al., 2023). Despite these advancements, the integration of explainable decision-making frame-works for UAS swarm coordination remains an underexplored topic. To assess the research landscape in this area, a literature review using the following terms: “explainable AI” and “UAS swarm” and “coordination” in the Scopus database has been done. The search was limited to publications from 2018 to 2025. From initial 127 results, only 14 paper addresses both UAS-swarm coordination and explainability aspects. Furthermore only 8 publications specifically examined explainable decision-making frameworks in the context of UAS-swarm operations. This confirms the significant research gap in integrating explainability into UAS-swarm coordination systems. Particularly those designed to support human operators.

## Explainable AI

XAI aims to bridge the gap between AI’s computational power and human trust by making AI systems more transparent and interpretable. The National Institute of Standards and Technology (NIST) identifies four core principles for XAI, which are listed below and are essential for fostering trust in AI systems deployed in safety-critical environments (Phillips et al., 2021).

- Transparency
- Interpretability
- Accuracy of explanations
- Knowledge limits

Human-centred XAI approaches emphasize tailoring explanations to the needs of specific user groups. For example, declarative explanations may suffice for engineers seeking technical details about system operations, while interactive or visual explanations may be more effective for non-technical users (Liao and Varshney, 2022). Studies have demonstrated that well-designed explanations can enhance user trust by providing clear justifications for AI decisions while reducing the need for constant monitoring (Ferrario and Loi, 2022; Friedrich et al., 2023).

In robotics and autonomous systems, XAI methods, such as decision trees have been employed to make AI-driven decisions more understandable to operators. These methods are particularly relevant for UAS swarms operating in dynamic environments, where real-time interpretability is crucial for effective human-machine collaboration (Ferrario and Loi, 2022).

## Integration of Reinforcement Learning With Fuzzy-Logic

RL has shown promise in optimizing decision-making systems by enabling adaptive learning from interactions with the environment (Kober et al., 2013; Sutton and Barto, 2018). However, traditional RL methods often lack transparency due to their reliance on complex neural network architectures. Integrating RL with Fuzzy-Logic offers a solution by combining the adaptability of RL with the interpretability of Fuzzy inference systems.

Takagi-Sugeno-Kang (TSK) Fuzzy-Systems optimized via RL have demonstrated success in various domains by fine-tuning membership functions and rule parameters without compromising system transparency (Zander et al., 2023; Berenji and Khedkar, 1998). For instance, RL can adjust Fuzzy-Rules dynamically based on feedback from simulation environments or real-world operations while ensuring that the underlying rule structure remains interpretable (Zander et al., 2023).

This hybrid approach has been applied to problems, such as mobile robot navigation and safety-critical decision-making in autonomous driving (Zander et al., 2023). By containing RL optimization within the bounds defined by the Fuzzy-Rules, it is possible to achieve a balance between performance improvement and explainability, which is a key requirement for trustworthy AI systems deployed in high-stakes environments.

## PROPOSED FRAMEWORK

### System Architecture

The in this paper proposed framework integrates the Fuzzy-Logic, RL and AI principles to create a decision support system for mission planning of UAS swarms. The architecture is designed to address two primary objectives. Firstly, the systems needs to ensure the interpretability and transparency through the use of Fuzzy-Logic and secondly it needs to be able to enhance the adaptability and performance via RL while maintaining explainability.

The system consists of three core components, which can be seen in Figure 1 and are described in the following section. The Fuzzy-Logic module defines linguistic variables, membership functions and rule-based inference systems to model human-like reasoning. It processes uncertain inputs and generates interpretable outputs in natural language. The RL module optimizes the parameters of the Fuzzy-Logic system, such as membership functions thresholds and rule weights. The optimization process is constrained to preserve the structure and interpretability of the fuzzy rules. The mission planning engine is the last core component, which integrates the outputs from the Fuzzy-Logic module with real-time data from UAS swarm operations to dynamically allocate tasks, manage resources and adapt to changes. The overall architecture ensures that decision-making remains transparent while leveraging adaptive learning capabilities to improve performance over time.

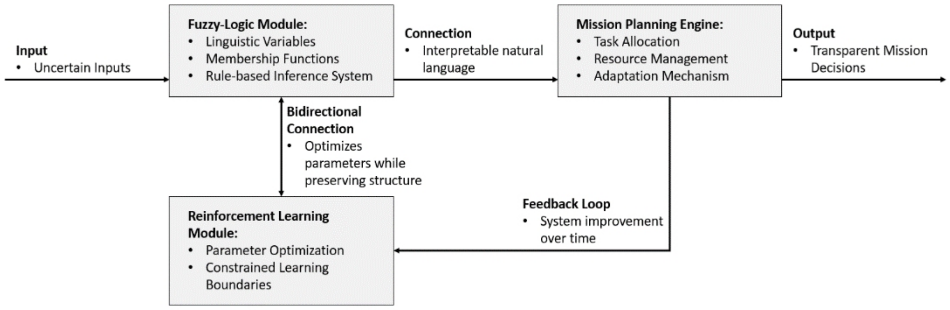


**Figure 1:** Three component architecture for explainable decision support UAS-swarm operations.

### Fuzzy-Logic Framework

In this framework, the linguistic variables are defined to represent key aspects of mission planning, such as risk level or resource availability. Each variable is associated with membership functions that map quantitative inputs, for example numerical sensor data, to qualitative terms, that are understandable by humans. These membership functions determine the degree to which an input value belongs to a specific linguistic category. For example, the extent to which a 70% battery level represents “sufficient” or “limited” resource availability.

A set of “if-then” rules forms the knowledge base for the decision-making process. For example, one of the Fuzzy-Rules could be defined as “If the risk level is high and the task priority is low, then postpone the current task”. Using these rules, which are designed in collaboration with domain experts, ensures the alignment with operational requirements and human reasoning patterns.

The Fuzzy-Logic framework employs a multi-step process. First is the input fuzzification, which converts precise numerical values into fuzzy-values. Then a rule evaluation in which each relevant rule is applied, is performed. After that an aggregation combines the output of all rules. And finally, defuzzification converts the aggregated fuzzy output into a clear, actionable decision or recommendation. This comprehensive process allows the system to handle uncertainties, while maintaining explainability throughout the decision process.

This approach differs from traditional AI systems as it preserves the reasoning process in a human-readable format, enabling operators to understand not only what decision was made, but also why it was made.

### Reinforcement Learning Integration

The RL component of the framework plays a crucial role in optimizing the fuzzy-logic system, while preserving the explainable nature. RL enhances the adaptability of the fuzzy-logic system by optimizing parameters based on feedback from simulations or real-world operations. It is then used to adjust the membership function boundaries to better reflect observed data distributions or it fine tunes the rule weights to prioritize certain actions under specific conditions.

To support transparency, the RL optimization is constrained with predefined boundaries. Therefore, RL cannot modify the structure or logic of the fuzzy-rules, but can adjust numerical parameters. This constraint-based approach implements a so called “cooperative neuro-fuzzy system”, where learning algorithms improve performance without sacrificing the interpretability advantages of the original fuzzy-system (Berenji and Khedkar, 1992). This hybrid approach ensures that the system remains interpretable while benefiting from adaptive learning capabilities.

The reward function needs to be carefully designed to balance performance metrics, such as mission completion time and resource efficiency, with explainable metrics, such as rule consistency and linguistic coherency. During the learning process, parameter changes are gradually introduced to allow for both system stability and human oversight of the adaptation process.

## Discussion: Human-Interpretable Re-Planning in Dynamic Environments

To evaluate the effectiveness of the proposed framework, simulation studies will be conducted using realistic mission scenarios, such as search and rescue operations in disaster areas involving UAS-swarms. The primary focus will be on examining how the system handles dynamic environmental changes, sensor failures and unexpected obstacles, which would typically require mission re-planning.

A critical aspect of the evaluation will focus on human interpretability. It will be evaluated, whether the operator better understands decision rationales from the fuzzy-logic enhanced system compared to traditional RL approaches. This comparative assessment builds on the fundaments of explainable AI, which includes interpretability metrics for AI system used in critical decision-making context (Arrieta et al., 2019).

The system's ability to re-plan missions when faced with environmental changes or asset failures represents a key evaluation topic, i.e. whether human operators can readily understand why re-planning occurred through the natural language explanations generated by the fuzzy-logic system. This human-centred evaluation approach aligns with current research on XAI, which emphasizes that the explanations must be adapted to human cognitive pat-terns rather than merely exposing technical processes (Miller, 2017).

By combining fuzzy-logic's inherent interpretability with RLs adaptability, it is expected to achieve a balance that maintains system transparency, while enabling performance optimization. Allowing operators to develop appropriate trust in the system's decision-making capabilities even as conditions change during mission execution.

## Discussion: A promising Approach and Way Forward

The integration of fuzzy-logic with RL for mission planning of UAS-swarms present a promising approach to address the critical challenges of maintaining human trust in increasingly autonomous systems. The following section discusses the anticipated benefits, evaluation approaches, challenges and broader implications of the proposed framework.

The primary advantage of the hybrid approach lies in its ability to balance adaptability with explainability, which is a crucial aspect in safety-critical environments. By preserving the interpretable structure of fuzzy-logic, while leveraging the optimization capabilities of RL, the system can evolve and improve without becoming opaque to human operators. This transparency is expected to foster appropriate trust from the operator.

Furthermore, the natural language explanations generated by the fuzzy-logic system offers justifications for decision that align with human cognitive patterns. Unlike black-box approaches that may provide no reasonable explanations, which do not accurately reflect the actual decision process, the proposed framework works with explanations which are directly emerged from the same rules to make decisions. This congruence between explanation and decision mechanism addresses a significant limitation in current explainable AI implementation for autonomous systems (Liao and Varshney, 2022; Arrieta et al., 2019).

The validation of the framework will follow a multi-phase approach focusing on both, system performance and human factors. Initial evaluation will occur though a simulation environment, which replicates challenging mission scenarios involving environmental changes, resource limitations and asset failures. These simulations will assess the system's ability to re-plan missions effectively while maintaining explainability. The evaluation will involve UAS operators with varying levels of experience to ensure the system's explanations are accessible across different expertise levels.

We hypothesize, that operators will demonstrate improved situation awareness, increased trust and more effective decision making when working with the explainable system compared to non-explainable alternatives. This human-centred evaluation approach acknowledges that the measurement of success for decision support systems is their ability to enhance human-AI collaboration rather than merely autonomous performance.

Furthermore, several challenges must be addressed during the implementation and evaluation of the proposed framework. First, the subjective nature of fuzzy-rule definition introduces potential inconsistencies or biases from domain experts. To mitigate this, it is planned to incorporate knowledge from multiple experts and refine rules though iterative validation.

Second, balancing the preservation of explainability with optimization through RL significantly presents a fundamental tension. If RL alters membership functions or rule weights, the resulting system may become less intuitive despite formal preservation of the rule structure. This needs to be addresses, to carefully constrain optimization spaces and ongoing evaluation of explanation quality throughout the learning process.

Lastly, the dynamic environments of disaster response scenarios present unpredictable challenges that may fall outside the anticipated parameters of the system. To address this limitation, the framework includes mechanisms for identifying situations, where confidence in recommendations should be reduced, explicitly communicating uncertainty to operators in these cases.

Beyond UAS-swarm operations, this research contributes to the growing field of human-AI teaming in complex environments. The principles and methodologies developed here may extend to other domains where explainable decision-making is critical, such as autonomous vehicles.

The approach represents a shift from viewing AI systems as either fully autonomous or merely tools, towards conceptualizing them as teammates with complementary capability to human operators. This perspective emphasizes the importance of mutual predictability, shared understanding and appropriate trust. This all is facilitated by the explainable AI framework.

As autonomous systems become increasingly prevalent in safety-critical domains, the ability to provide transparent justifications for decisions will likely become a regulatory requirement rather than merely a design preference. This framework anticipates such a development by demonstrating how explainability can be preserved even as systems learn and adapt over time.

## REFERENCES


Arrieta A. B., Diaz-Rodriguez N., Del Ser J. and Herrera F. (2019) Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges towards responsible AI. Elsevier. http://dx.doi.org/10.48550/arXiv.1910.10045

Berenji H. R. and Khedkar P. (1992) Learning and tuning fuzzy-logic controllers through reinforcements. IEEE. https://doi.org/10.1109/72.159061

Berenji H. R. and Khedkar P. S. (1998) Usig fuzzy-logic for performance evaluation in reinforcement learning. NASA. https://doi.org/10.1016/S0888-613X(97)10007-X

Crandall J. W. and Cummings M. L. (2007) Developing performance metrics for the supervisory control of multiple robots. ACM. https://doi.org/10.1145/1228716.1228722

Endsley M. R. (2016) From here to autonomy: Lessons learned from human-automation research. Human Factors. https://doi.org/10.1177/0018720816681350

Erdelj M., Krol M. and Natalizio E. (2017) Wireless sensor networks and multi-UAV systems for natural disaster management. Computer Networks. http://dx.doi.org/10.1016/j.comnet.2017.05.021

Ferrario A. and Loi M. (2022) How explainability contributes to trust in AI. ACM. https://doi.org/10.1145/3531146.3533202

Friedrich M. (2021) Human machine interface design to support safety risk monitoring of autonomous small unmanned aircraft systems – Results from a mock-up evaluation. IEEE. https://doi.org/10.1109/DASC52595.2021.9594511

Friedrich M., Küls J., Findeisen M. and Peinecke N. (2023) HMI design for explainable machine learning enhanced risk detection in low-altitude UAV operations. IEEE. https://doi.org/10.1109/DASC58513.2023.10311252

Huttner J.-P. and Friedrich M. (2023) Current challenges in mission planning systems for UAVs: A systematic review. ICNS. https://doi.org/10.1109/ICNS58246.2023.10124299

Improta G., Mazzella V., Vecchione D., Santini S. and Triassi M. (2019) Fuzzy-logic based clinical decision support system for the evaluation of renal function in post-transplant patients. Wiley. https://doi.org/10.1111/jep.13302

Karampelia I., Kyriakidis T. and Louta M. (2023) UAV swarms & task allocation: The way ahead in precision agriculture. IEEE. https://doi.org/10.1109/IISA59645.2023.10345854

Kober J., Bagnel J. A. and Peters J. (2013) Reinforcement learning in robotics: A survey. The International Journal of Robotics Research. http://dx.doi.org/10.1177/0278364913495721

Lakhmissi C. and Boumehraz M. (2013) Fuzzy-Logic and reinforcement learning based approaches for mobile robot navigation in unknown environment. University of Biskra. http://archives.univ-biskra.dz/handle/123456789/7511

Lamont G. B., Slear J. and Melendez K. (2007) UAV swarm mission planning and routing using multi-objective evolutionary algorithms. IEEE. https://doi.org/10.1109/MCDM.2007.369410

Li Y., Wu B., Huang Y. and Luan S. (2024) Developing trustworthy artificial intelligence: Insights from research on interpersonal, human-automation, and human-AI trust. Front. Psychol. 15:1382693. https://doi.org/10.3389/fpsyg.2024.1382693

Liao Q. V. and Varshney K. R. (2022) Human-centred explainable AI: From algorithms to user experiences. IBM Research. https://doi.org/10.48550/arXiv.2110.10790

Mehrotra S., Degachi C., Vereschak O., Jonker C. M. and Tielman M. L. (2023) A systematic review on fostering appropriate trust in human-AI interaction. Cornell University. https://doi.org/10.48550/arXiv.2311.06305

Melin P. and Castillo O. (2013) A review on the applications of type-2 fuzzy-logic in classification and pattern recognition. ScienceDirect. http://dx.doi.org/10.1016/j.eswa.2013.03.020

Miller T. (2017) Explanation in artificial intelligence: Insights from the social sciences. Elsevier. http://dx.doi.org/10.1016/j.artint.2018.07.007

Peng Q., Wu H. and Li N. (2022) Modeling and solving the dynamic task allocation problem of heterogeneous UAV swarm in unknown environment. Hindawi. https://doi.org/10.1155/2022/9219805

Peng Q., Wu H. and Xue R. (2021) Review of dynamic task allocation methods for UAV swarms oriented to ground targets. ISSN. https://doi.org/10.23919/CSMS.2021.0022

Phillips P. J., Hahn C. A., Fontana P. C., Yates A. N., Greene K., Broniatowksi D. A. and Przybocki M. A. (2021) Four principles of explainable artificial intelligence. NIST. https://doi.org/10.6028/NIST. IR.8312

Polka M., Ptak S. and Kuziora L. (2017) The use of UAVs for search and rescue operations. Procedia Engineering. http://dx.doi.org/10.1016/j.proeng.2017.06.129

Seeber I., Bittner E., Briggs R. O. and Söllner M. (2018) Machines as teammates: A collaboration research agenda. Universität Hamburg. http://dx.doi.org/10.24251/HICSS.2018.055

Sutton R. S. and Barto A. G. (2018) Reinforcement learning: An introduction. MIT Press. ISBN 978–0262039246

Tahir A., Böling J. M., Haghbayan H. and Plosila J. (2019) Swarms of unmanned aerial vehicles – A survey. Elsevier. http://dx.doi.org/10.1016/j.jii.2019.100106

Thiebes, S., Lins, S. & Sunyaev, A. (2021) Trustworthy artificial intelligence. Electron Markets 31, 447–464. https://doi.org/10.1007/s12525-020-00441-4

Wu H. and Xu Z. S. (2021) Fuzzy-Logic in decision support: Methods, applications and future trends. Sichuan University. https://doi.org/10.15837/ijccc.2021.1.4044

Yan Y., Sun Z., Hou Y., Zhang B., Yuan Z., Zhang G., Wang B. and Ma X. (2023) UAV swarm mission planning and load sensitivity analysis based on clustering and optimization algorithms. Appl. Science. https://doi.org/10.3390/app132212438

Zander E., van Oostendorp B. and Bede B. (2023) Reinforcement learning with takagi-sugeno-kang fuzzy systems. Compley Eng. Systems. https://doi.org/10.20517/ces.2023.11