UDC 65.012:656
DOI https://doi.org/10.32782/2663-5941/2026.3.1/27

Palonyi A.S.
https://orcid.org/0000-0003-1539-3671
Ukrainian State Flight Academy

Nechypurenko A.G.
https://orcid.org/0009-0002-1691-4424
Ukrainian State Flight Academy

OPTIMIZATION OF AUTOMATED STRESS MANAGEMENT STRATEGIES FOR AIR TRAFFIC CONTROLLERS USING REINFORCEMENT LEARNING METHOD

The subject matter of the article is the occupational stress of air traffic controllers (ATCOs) in the context of increasing air traffic intensity and the growing complexity of the professional environment, as well as the mechanisms for proactive management of work organization parameters. The research focuses on the conceptualization and mathematical modeling of intelligent agents based on deep reinforcement learning (RL) methods within a hybrid decision support system (DSS) designed to maintain the psychophysiological state of the air traffic controller within the functional optimum (eustress). The aim of the article is the development and functional modeling of RL agents for adaptive stress management and the continuous optimization of human-centric support strategies. The tasks of the article include: formalization of the interaction between the operator, the environment, and the agent; implementation of an intermediate state abstraction layer; application of the "action stacking" concept to structure tactical, operational, and strategic interventions; and synthesis of a multi-objective reward function with hybrid expert tuning. Research methods include the Partially Observable Markov Decision Process (POMDP), reinforcement learning (RL), inverse reinforcement learning (IRL), fuzzy logic as a cognitive interface, and explainable artificial intelligence (XAI) methods. The obtained results include: a formalized POMDP model with a three-component observation vector; an integrated fuzzy abstraction layer that transforms raw multimodal data into a belief vector; an ordered matrix of batched interventions; and an extended multi-objective reward function that minimizes the risks of hypostimulation and distress. The scientific novelty lies in the integration of a hierarchical fuzzy inference system with RL algorithms as a deterministic state abstraction layer, as well as in proposing a hybrid reward function architecture that combines classical RL with inverse learning to extract a hidden expert function from the historical decision trajectories of experienced shift supervisors. The practical significance of the system lies in facilitating a technological transition from reactive monitoring to proactive real-time stress management, where transparent recommendations regarding resectorization, micro-breaks, and automation level adjustments are provided to supervisors using XAI tools (Solution Space Diagram).

Keywords: *air traffic controller, occupational stress, decision support system, reinforcement learning, explainable artificial intelligence.*

Formulation of the problem. The current stage of development of the aviation industry is characterized by an increase in air traffic intensity, the growing complexity of airspace structure, the integration of unmanned systems, and the development of advanced air mobility. The EUROCONTROL Seven-Year Forecast 2025-2031 states that by 2031, approximately 12.2 million flights are expected with an average annual growth rate of about 2.0%, while the long-term forecast 2024–2050 envisions a baseline scenario of 15.4 million flights in 2050 [1, 2]. The quantitative growth of air traffic is accompanied by a deterioration in operational performance. According to the Performance Review Report (PRR) 2024 (EUROCONTROL), the total volume of en-route ATFM delays in 2024 amounted to 22.4 million minutes (equivalent to 2.8 billion Euro delay costs), which is approximately 24% higher than in 2023 (en-route phase) and is the worst indicator since 2001. Such dynamics indicate not only system overload but also an increase in cognitive and organizational pressure on air traffic control (ATC) personnel [3]. A significant portion of these delays is associated precisely with staffing and sector capacity constraints in several congested nodes of the European airspace.

The operational environment is further complicated by external technological threats. PRR 2024 recorded that up to 38% of European en-route traffic passed through regions that periodically or regularly experienced GNSS radio frequency interference (GNSS RFI). This implies an increased need for alternative navigation procedures, radar vectoring, and additional actions by the air traffic controller (ATCO), consequently leading to an increase in cognitive workload during real-time sector management. In addition, the safe integration of new types of airborne operations requires the restructuring of airspace, ATC procedures, and support tools for ATCOs. In turn, SESAR considers such changes as a factor driving the growth of operational environment complexity and personnel workload [4].

EASA results showed that in the study sample, 5.6% of all analyzed controller shifts were associated with critical fatigue. At the same time, over the 2013–2022 period in EASA Member States and the United Kingdom, no accidents or serious incidents were directly classified as caused by ATCO fatigue, although 184 occurrences were identified in European event databases where ATCO fatigue was mentioned as a contributing factor; of these, 59 had a potentially hazardous outcome, including loss of separation, airprox, and the issuance of incorrect instructions [5]. The greatest practical value of the EASA report lies in the quantitative assessment of risk factors presented in Fig. 1.

In the USA, similar conclusions were formulated in the FAA expert panel report “Assessing Fatigue Risk in FAA Air Traffic Operations” (2024), which contains 58 opportunities/recommendations regarding fatigue risk mitigation. Experts explicitly recommend abandoning 2-2-1 type rotational schedules, acknowledging them as physiologically harmful due to circadian disruption. Special emphasis is placed on the necessity to provide sufficient off-duty time, in particular at least 10–12 hours between shifts, especially before night duties [6].

The aviation industry's response to the challenges of fatigue was the development and implementation of Fatigue Risk Management Systems (FRMS). ICAO in Doc 9966 defines a framework for the oversight and implementation of fatigue risk management approaches based on scientific data, operational information, and continuous monitoring. Within the European legal framework, issues of licensing, competence, and general requirements for ANS providers are regulated, in particular, by Commission Regulation (EU) No 2015/340 and Commission Implementing Regulation (EU) No 2017/373 [7-9]. However, EASA research shows that even with the formal existence of FRMS policies, their informational sensitivity is often limited. About 40% of providers collected data on sleep and fatigue only quarterly or annually [5]. Traditional FRMS primarily operate as reactive-planning mechanisms: they are well-suited for schedule verification, event statistics tracking, and regulatory control, but perform significantly worse at detecting short-term surges of cognitive overload that occur directly during tactical operations in the ATC sector. It is this methodological limitation that justifies the transition to systems capable of utilizing multimodal data streams and supporting real-time management decisions.

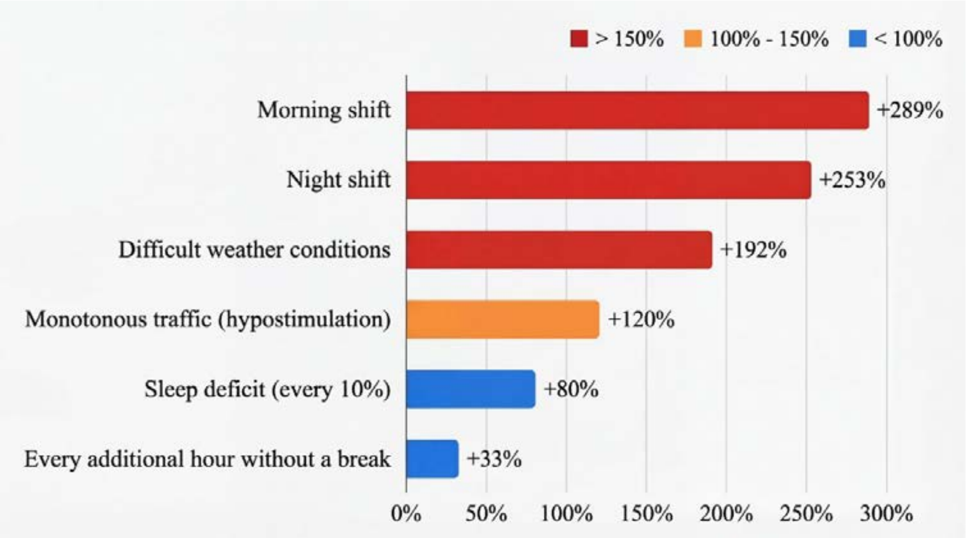

Fig. 1. Factors increasing the risk of critical fatigue in ATCOs

Analysis of recent research and publications. The air traffic management (ATM) modernization paradigm implemented under the SESAR and NextGen programs is aimed at a deeper utilization of automation, data analytics, and artificial intelligence (AI). However, Lisanne Bainbridge, in her classic work "Ironies of Automation," demonstrated that excessive or opaque automation does not eliminate the human factor, but may alter the nature of operator errors, reduce their engagement, and create new risks during moments of system failure or unpredictable behavior [10]. In contemporary ATM research, this thesis has transformed into the requirement to design not just automated, but human-centric decision support systems (DSS). Intelligent decision support complexes in hybrid "human-AI" systems must adapt to the context, workload, and state of the ATCO. Such an approach is part of the strategic directions of the SESAR program and serves as the foundation for research projects like AWARE, in which the AI assistant is considered an adaptive, human-centric component of ATCO support, rather than a complete replacement for the human [11-13]. In this context, it is both scientifically and practically justified to employ an approach where operator state assessment systems (specifically fuzzy models and multi-factor inference systems) are combined with reinforcement learning (RL) algorithms for dynamic, real-time optimization of workload management strategies. Such hybrid DSSs should integrate traffic parameters, weather conditions, sector context, staffing constraints, type of controller shift, automation mode, degree of human involvement in the control loop, circadian factors, and, if possible, current indicators of the ATCO's functional state for proactive workload adaptation. In such a setup, the RL algorithm acts as an adaptive mechanism for finding optimal strategies for rotation, breaks, sector redistribution, and forms of ATCO assistance in a specific operational context. A relevant advanced development is the CODA (Controller adaptive Digital Assistant) project, aimed at creating a digital assistant for ATCOs. The system predicts future traffic and assesses the ATCO's psychophysiological state – mental workload, attention, stress, fatigue, vigilance, and the capacity to handle expected workload. On this basis, CODA adapts the level of automation: it can enhance automated support, activate AI-based tools, or initiate changes in airspace organization, specifically sector splitting. Technologically, the project combines prediction models, neurophysiological assessment, and adaptive automation for dynamic task allocation between the ATCO and the digital assistant based on cognitive complexity; the optimization of such interaction is carried out within the Human–Machine Performance Envelope approach [14].

In order for intelligent algorithms to effectively optimize stress management strategies, the system requires a clear mathematical and logical model of the control object, i.e., a model of the formation of occupational stress itself. In our previous research on the system dynamics of ATCO stress [15], a comprehensive classification of occupational stress sources was proposed (32 key specific ATCO stressors structured into six conceptual clusters by their source: operational, ergonomic, informational-cognitive, social, psychological, and organizational). Each of these factors is assigned specific indices: an index of controllability and an index of duration of exposure, which allows differentiating measures into operational, tactical, and strategic. Based on this clustering, a multi-criteria system dynamics model of occupational stress has been built. According to this model, the impact of distributed stressors is accumulated and transformed through five macro-components – feedback aggregators: the level of subjective sense of control over the work situation (PC), the level of perceived quality of teamwork (TW), the level of subjective perception of current job demands and complexity (JD), the level of accumulated background stress (BS), and the level of stress self-management efficacy (SM). A specific feature of this model is the identification of non-linear feedback loops that account for a self-reinforcing effect: an objective increase in traffic complexity elevates the overall stress level, which, in turn, physiologically narrows the cognitive capacity of the brain, causing those same work demands to seem even more insurmountable to the ATCO. Similarly, an increase in stress reduces the sense of control over the situation, provoking further panic-driven stress escalation. It is exactly this cyclic nature of occupational stress development that demands intelligent proactive and reactive support for real-time ATCO stress management. For the technological implementation of this system dynamics of ATCO occupational stress, within the framework of our previous studies, the concept of a hybrid decision support system for adaptive real-time assessment of ATCO stress with the capability to generate personalized recommendations for management interventions was proposed [16]. Proactive dynamic management of ATCO occupational stress requires the involvement of the mathematical apparatus of reinforcement learning.

Task statement. The main goal of the article is the development and modeling of the functioning of intelligent agents based on deep reinforcement learning methods, as part of a hybrid DSS for adaptive management of organizational environment parameters and ensuring the ATCO's psychophysiological state remains within the functional optimum.

Outline of the main material of the study. The implementation of a DSS in the format of a proactive digital manager of ATCO working conditions requires a formal description of the interaction process between the controller, the environment, and the intelligent agent. The most adequate mathematical framework for this is the Partially Observable Markov Decision Process (POMDP). This choice is determined by the fact that the actual psychophysiological and cognitive state of the ATCO cannot be directly measured with absolute accuracy; the system deals only with indirect multimodal features, based on which the operator's hidden state is estimated.

The process is defined by an extended tuple:

$$M_{POMDP} = \langle S, A, P, R, \Omega, O, \gamma \rangle \tag{1}$$

where *S* is the set of hidden states of the ATCO and the environment; *A* is the set of admissible control actions; *P* is the state transition function; *R* is the reward function; *Ω* is the set of possible observations; *O* is the observation function; *γ* is the discount factor for future rewards.

Since the true state of the ATCO is hidden, the agent does not operate directly with the state $s_t \in S$, but operates with an observation vector $o_t \in \Omega$, formed at each time step Δt. Such a vector is represented as a three-component structure:

$$o_t = \left(O_{traffic(t)}, O_{physio(t)}, O_{context(t)} \right) \tag{2}$$

where $O_{traffic(t)}$ is a block of objective air traffic situation parameters, including the number of aircraft in the sector, the geometric complexity index of the airspace configuration, traffic density, and the frequency of potentially conflicting convergences; $O_{physio(t)}$ is a block of physiological indicators, such as heart rate variability, electrodermal activity, respiration rate, and other sensory markers of stress; $O_{context(t)}$ is a contextual block that includes the phase of the controller shift, the duration of continuous work, accumulated sleep deficit, circadian phase, and other relevant conditions.

The state transition function (P) formalizes the dynamics of the system state in the POMDP model, i.e., it specifies how the current hidden state changes after the agent's action:

$$P\left(s_{(t+1)} \mid s_t, a_t \right) = Pr\left(s_{(t+1)} \mid s_t, a_t \right) \tag{3}$$

where s_t is the current hidden state of the ATCO and the environment at time t (fatigue level, cognitive workload, and stress level of the ATCO, traffic complexity); a_t is the action of the agent at time t (changing the level of automation, resectorization, recommended micro-breaks, etc.); $s_t + 1$ is the next hidden state of the system after the execution of action a_t, i.e., the state at time t+1.

The observation function (O) formalizes the mechanism for generating a new feature vector (sensory and contextual data) from the actual hidden system state and the executed action:

$$O\left(o_{(t+1)} \mid s_{(t+1)}, a_t \right) = Pr\left(o_{(t+1)} \mid s_{(t+1)}, a_t \right) \tag{4}$$

where $o_t + 1$ is the observation at time t+1; a_t is the agent's action performed at the previous step (the action can affect not only the state but also what data will be available for observation).

A key feature of the proposed approach is the use of an already developed hierarchical fuzzy inference system as a deterministic state abstraction layer [16]. Such a layer transforms the high-dimensional observation vector into a belief vector b_t , suitable for further processing by the RL agent:

$$b_t = FIS\left(o_t \right) = \left(PC_t, JD_t, TW_t, BS_t, SM_t, S_{overall}\left(t \right) \right) \tag{5}$$

where $PC_t, JD_t, TW_t, BS_t, SM_t$ are the respective variables represent aggregated fuzzy indicators of the occupational stress assessment components, a $S_{overall}$ is the integral metric indicates the current overall stress level ($S_{overall}(t) \in [1, 7]$).

Thus, fuzzy logic acts as an intermediate cognitive interface between raw physiological and operational data and the reinforcement learning agent.

Taking into account the concept of "action-stacking", an individual agent action is not an elementary command, but an ordered package of interconnected interventions:

$$a_t = \left(a_t^{(tac)}, a_t^{(opr)}, a_t^{(str)} \right) \tag{6}$$

where $a_t^{(tac)}$ are ***tactical actions*** are related to adaptive changes in the level of automation and the automatic connection of auxiliary AI modules (on non-critical tasks) to unload the ATCO, local redistribution of tasks within the shift, introduction of short micro-breaks, and transferring the ATCO to a less stressful workplace; $a_t^{(opr)}$ are ***operational actions*** are related to opening or closing sectors, changing sector configurations based on projected demand, providing additional team support, redistributing workload among sectors/positions, activating reserve positions, correcting the break plan within the shift, and implementing local air traffic flow and capacity management measures; $a_t^{(str)}$ are ***strategic actions*** are related to rostering and duty policies, regulating minimum rest time between shifts, imposing limits on consecutive night shifts and overtime, revising the ratio of time on position to breaks, recalculating the minimum number of personnel per shift, providing cross-rating/training procedures for flexible sector opening, and designing the sector system and pre-prepared configurations.

Thus, the RL agent does not merely react to a state change but forms a structured scenario of organizational impact on the ATCO's working environment.

A critical problem of the model is formulating the reward function. In the stress management task, the target state is not minimizing stress to zero, but maintaining it within the optimal functional tone, i.e., in the ***zone of eustress***. According to the provisions of the Yerkes-Dodson law, both excessive stress and insufficient stimulation can degrade performance quality. For instance, hypostimulation (very low traffic levels) reduces ATCOs' vigilance and leads to errors.

We utilize a generalized multi-objective reward function:

$$R_t = \omega_1 R_{eustress}\left(t\right) + \omega_2 R_{perf}\left(t\right) + \omega_3 R_{cost}\left(a_t\right) + \omega_4 R_{safety}\left(t\right) \quad (7)$$

where $\omega_1, \omega_2, \omega_3, \omega_4$ – where the weights are the coefficients of individual target components; $R_{eustress}$ – the eustress component is defined as an inverted quadratic function, peaking at the point of optimal workload: $R_{eustress}\left(t\right) = 100 - \lambda \cdot (S_{overall}(t) - S_{optimal})^2$, $S_{optimal}$=3.5; $R_{perf}\left(t\right)$ is the efficiency component acts as a reward for maintaining sector capacity without delays; $R_{cost}\left(a_t\right)$ is the action cost represents the operational "cost" of the intervention (e.g., a penalty for frequent sector configuration changes); $R_{safety}(t)$ is a penalty that grows exponentially if the overall stress exceeds the critical limit of 5.0 (distress zone and loss of situational awareness) (Fig. 2).

The safety component must have a strictly asymmetrical nature, as transitioning into the distress zone is significantly more dangerous than briefly dropping below the optimum. Therefore, the penalty for exceeding the critical threshold is defined by an exponential dependency:

$$R_{safety}(t) = \left\{ \begin{array}{l} 0, \text{ if } S_{overall}(t) \le 5.0; \\ -\eta \cdot \left(e^{\kappa \cdot (S_{overall}(t) - 5.0)} - 1 \right), \text{ if } S_{overall}(t) > 5.0 \end{array} \right\} \quad (8)$$

The action cost component is used to prevent the system from generating an excessive number of reorganizational decisions:

$$R_{cost}(a_t) = -c(a_t) - \rho \cdot 1(a_t \ne a_{t-1}) \quad (9)$$

where $c(a_t)$ is the first part acts as the base operational cost of the selected measure, and the second term implements an action damping mechanism, i.e., penalizes frequent changes in management decisions.

Manual tuning of the weight coefficients ($\omega_1 \ldots \omega_4$) in the reward function is an extremely complex task, as human decision-making in civil aviation involves subtle judgments, risk tolerance, and adaptive strategies. To solve this problem, a hybrid application of traditional RL and Inverse Reinforcement Learning (IRL) is proposed. Thus, we obtain a learning architecture based on expert demonstrations with subsequent optimization.

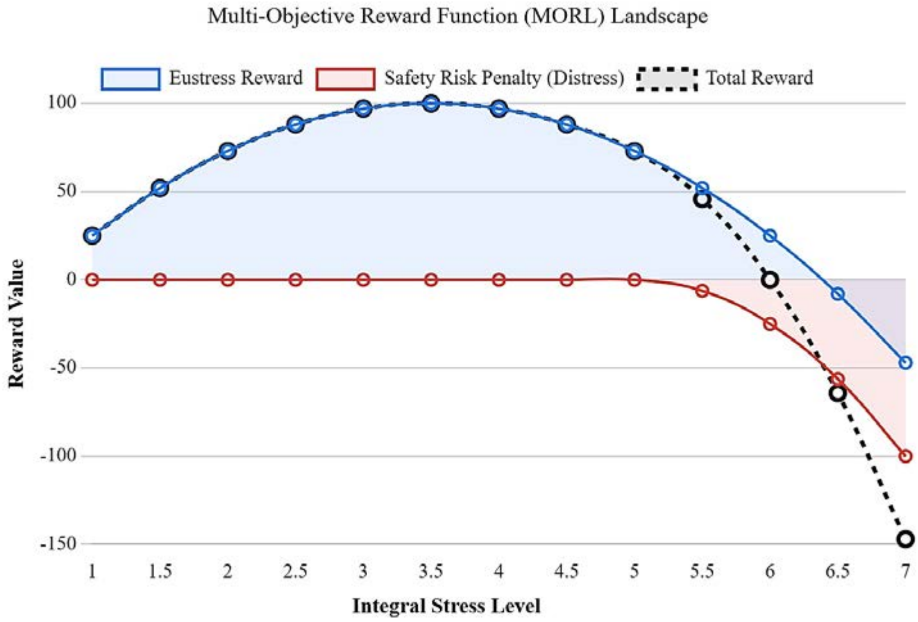

Fig. 2. Dependence of the RL-agent reward function components on the integral stress level of the air traffic controller

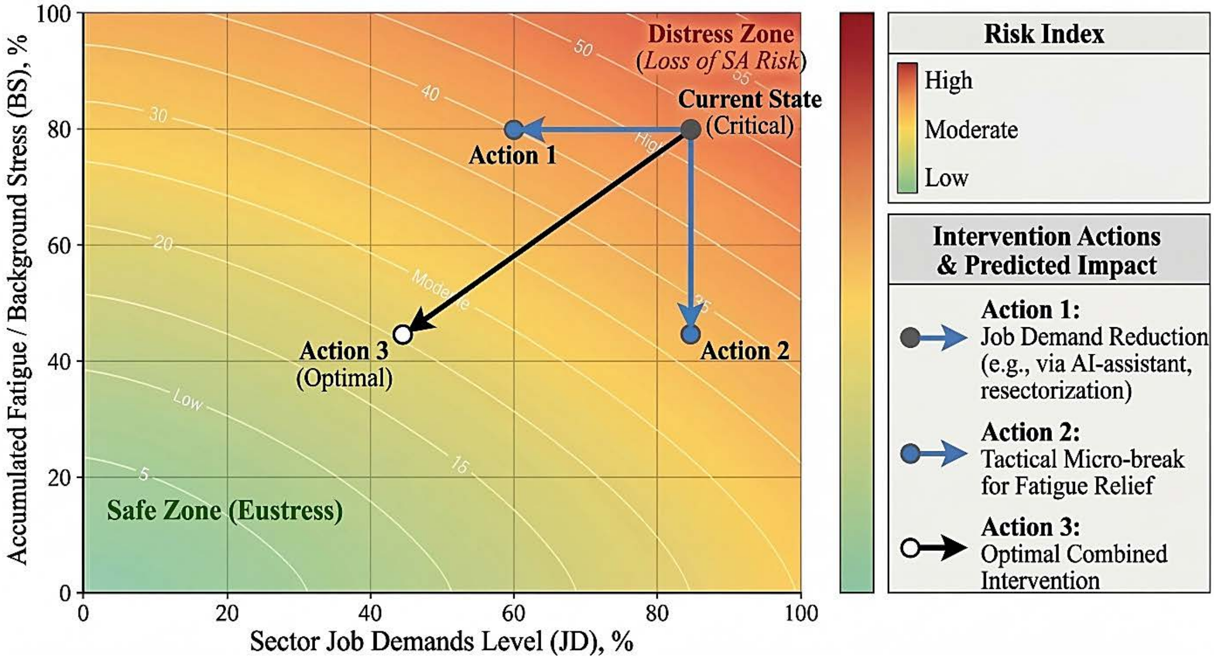

Fig. 3. Solution space visualization (XAI) for the shift supervisor

In the first stage, a hidden expert reward function is extracted from large datasets of historical data:

$$R_{IRL} = f\left(\tau_{expert}\right) \tag{10}$$

where τ_{expert} is the corresponding variable denotes the action trajectories of experienced supervisors and ATCOs in critical and pre-critical workload regimes (the system analyzes the actions of the most experienced supervisors during peak loads: exactly when they decided to split a sector, when they sent an ATCO on a break, and when they allowed them to work at the edge of eustress).

In the second stage, this function is combined with explicitly defined safety rules within a hybrid objective function:

$$R_{hybrid}(s_t, a_t) = \alpha \cdot R_{explicit}(s_t, a_t) + \beta \cdot R_{IRL}(s_t, a_t) \tag{11}$$

where $R_{explicit}$ is the strict normative and technological constraints that the system has no right to violate under any circumstances; R_{IRL} is the empathetic component derived from data, which stimulates decision-making in the style of the best human practices; α and β are balancing coefficients.

After constructing the hybrid reward function, a classical safe reinforcement learning algorithm searches for the optimal control policy:

$$\pi^* = \arg\max_{\pi} E_{\pi}\left[\sum_{t=0}^{T} \gamma^t R_{hybrid}\left(s_t, a_t\right)\right] \tag{12}$$

where π^* is the optimal policy; E_{π} is the mathematical expectation over trajectories generated by the policy; T is the planning horizon; γ^t is the discount factor, which determines the relative value of delayed consequences.

Training of such an agent is carried out in a simulator environment, where rare scenarios of peak workload, chronic fatigue, night shifts, and combined human-machine failures can be reproduced without risk to flight safety.

To ensure transparency of the system's operation and trust from supervisors and ATCOs, Explainable AI (XAI) is integrated into the proposed architecture. Every management action of the agent is accompanied by an interpretable explanation that links the current integral stress index, the most significant stressors, and the expected effect of the recommended intervention. In this context, it is advisable to use a Solution Space Diagram (SSD), which visualizes why, given a specific value of the state, sector workload level, and accumulated fatigue, the system chose automation adaptation, resectorization, a micro-break, or another measure (Fig. 3).

Conclusions. The proposed model enables the implementation of a closed-loop intelligent control system for the ATCO's working conditions. In the considered system, multimodal observations are transformed by a fuzzy system into a compact belief vector; on this basis, the RL agent assesses the current risk of exiting the eustress state, selects a hierarchical control action, predicts its consequences for safety and capacity, and then explains its recommendation to the operator in an interpretable format. Such an architecture lays the groundwork for transitioning from reactive fatigue monitoring to proactive, adaptive, and human-centric stress management in the ATC environment.

An important stage for further research is the validation of the hybrid decision support system for managing ATCO' occupational stress on representative datasets of real operational shifts. It is necessary to develop methods for adaptive dynamic adjustment of the threshold values of ATCOs' occupational stress. Such models should take into account not only general indicators but also the individual ATCO's experience level, their sensitivity to specific stressors, and their capacity for occupational stress self-management. This entails creating algorithms that continuously update the stress resilience profiles of each specialist, allowing the system to propose highly personalized interventions. Furthermore, an important task is the application of deep learning methods for the automatic identification of new stressors, their cumulative effects, stress response patterns, and adaptive proactive management interventions that account for the organizational and technological changes occurring in the modern ATM system.

Bibliography:

1. EUROCONTROL. Seven-Year Forecast 2025–2031. Brussels: EUROCONTROL, Spring 2025. 23 p. URL: https://www.eurocontrol.int/sites/default/files/2025-02/eurocontrol-seven-year-forecast-2025-2031-february-2025.pdf (accessed: 16.03.2026).
2. EUROCONTROL. Aviation Long-Term Outlook: Flights and CO_2 emissions forecast 2024-2050. Brussels: EUROCONTROL, 2024. 40 p. URL: https://www.eurocontrol.int/sites/default/files/2024-12/eurocontrol-aviation-long-term-outlook-flights-co2-forecast-2024-2050.pdf (accessed: 16.03.2026).
3. PRR 2024. Performance Review Report. An Assessment of Air Traffic Management in Europe 2024. Brussels: EUROCONTROL, Performance Review Commission, March 2025. 106 p. URL: https://www.eurocontrol.int/sites/default/files/2025-03/eurocontrol-performance-review-report-2024.pdf (accessed: 16.03.2026).
4. Federal Aviation Administration. Advanced Air Mobility (AAM) Implementation Plan. Version 1.0. Washington, DC : FAA, 2023. 40 p.; SESAR Joint Undertaking. European ATM Master Plan. 2025 edition. Brussels : SESAR Joint Undertaking, 2025. 90 p.
5. European Union Aviation Safety Agency. Study on the Analysis, Prevention and Management of Air Traffic Controller Fatigue. Cologne : EASA, May 2024. 138 p. URL: https://www.easa.europa.eu/sites/default/files/dfu/ATCO_Fatigue_Study_Final_report_D4C1_-_V4.1_final_21_May_2024.pdf (accessed: 16.03.2026).
6. Federal Aviation Administration. Assessing Fatigue Risk in FAA Air Traffic Operations. Washington, DC: FAA, 2024. 114 p. URL: https://www.faa.gov/newsroom/media/Fatigue_Report.pdf (accessed: 16.03.2026).
7. ICAO Doc 9966. Manual for the Oversight of Fatigue Management Approaches. Montréal: ICAO, 2016. 202 p.
8. Commission Implementing Regulation (EU) 2017/373 of 1 March 2017 laying down common requirements for providers of air traffic management/air navigation services and other air traffic management network functions and their oversight. Official Journal of the European Union. L62, 8.03.2017. 126 p.
9. Commission Regulation (EU) 2015/340 of 20 February 2015 laying down technical requirements and administrative procedures relating to air traffic controllers’ licences and certificates. Official Journal of the European Union. L63, 6.03.2015. 122 p.
10. Bainbridge L. Ironies of automation // Automatica. 1983. Vol. 19, iss. 6. P. 775–779. DOI: 10.1016/0005-1098(83)90046-8.
11. SESAR Joint Undertaking. European ATM Master Plan. 2025 edition. Brussels : SESAR Joint Undertaking, 2025. 90 p.
12. AWARE project (SESAR Joint Undertaking): Achieving human-machine collaboration with artificial situational awareness : website. URL: https://www.aware-sesar.eu (accessed: 16.03.2026).
13. Pham D.-T. et al. Human-AI Hybrid Paradigm for Collaborative Air Traffic Management Systems. SESAR Innovation Days, Rome, 12-15 November 2024. 9 p. URL: https://www.sesarju.eu/sites/default/files/documents/sid/2024/papers/SIDs_2024_paper_087%20final.pdf (accessed: 16.03.2026).
14. Periodic Reporting for period 1 – CODA (COntroller adaptive Digital Assistant). CORDIS : European Commission. 4 p. DOI: https://doi.org/10.3030/101114765.
15. Palonyi A.S., Nechypurenko A.G. Modelling the dynamics of occupational stress in air traffic control. Вчені записки таврійського національного університету імені В.І. Вернадського. Серія: Технічні науки. 2024. Том 35 (74), №5. P. 231-239. DOI: https://doi.org/10.32782/2663-5941/2024.5.1/34.
16. Palonyi A.S., Nechypurenko A.G. Hybrid decision support system for adaptive assessment and management of air traffic controllers’ occupational stress in real-time. Вчені записки таврійського національного університету імені В.І. Вернадського. Серія: Технічні науки. 2025. Том 36 (75), №6. С. 291-299. DOI: https://doi.org/10.32782/2663-5941/2025.6.2/40.

Пальоний А.С., Нечипуренко А.Г. ОПТИМІЗАЦІЯ СТРАТЕГІЙ АВТОМАТИЗОВАНОГО СТРЕС-МЕНЕДЖМЕНТУ АВІАДИСПЕЧЕРІВ МЕТОДАМИ НАВЧАННЯ З ПІДКРІПЛЕННЯМ

Предметом статті є професійний стрес авіадиспетчерів в умовах зростання інтенсивності повітряного руху та ускладнення професійного середовища, а також механізми проактивного управління параметрами організації праці. Дослідження зосереджено на концептуалізації та математичному моделюванні інтелектуальних агентів на основі методів глибокого навчання з підкріпленням (RL) у складі гібридної системи підтримки прийняття рішень (СППР) для утримання психофізіологічного стану авіадиспетчера в межах функціонального оптимуму (еустресу). Метою статті є розробка та моделювання функціонування RL-агентів для адаптивного управління стресом та безперервної оптимізації стратегій людиноцентричної підтримки. Завдання статті включають: формалізацію взаємодії між оператором, середовищем та агентом; реалізацію проміжного шару абстракції стану; застосування концепції «комбінування заходів» для структурування тактичних, оперативних і стратегічних втручань; синтез багатокритеріальної функції винагороди з гібридним експертним налаштуванням. Методи дослідження включають частково спостережуваний марковський процес прийняття рішень (POMDP), навчання з підкріпленням (RL), обернене навчання з підкріпленням (IRL), нечітку логіку як когнітивний інтерфейс та методи зрозумілого штучного інтелекту (XAI). Отримані результати охоплюють: формалізовану модель POMDP із трикомпонентним вектором спостережень; інтегрований нечіткий шар абстракції, що перетворює сирі мультимодальні дані на вектор переконань; впорядковану матрицю пакетних втручань; розгорнуту багатокритеріальну функцію винагороди, що мінімізує ризики гіпостимуляції та дистресу. Наукова новизна полягає в інтеграції ієрархічної нечіткої системи виводу з RL-алгоритмами як детермінованого шару абстракції стану, а також у пропозиції гібридної архітектури функції винагороди, що поєднує класичне RL з оберненим навчанням для вилучення прихованої експертної функції з історичних траєкторій рішень досвідчених супервайзерів диспетчерських змін. Практична значимість системи полягає у забезпеченні технологічного переходу від реактивного моніторингу до проактивного стрес-менеджменту в реальному часі, де за допомогою інструментів XAI (діаграма простору рішень) супервайзерам надаються прозорі рекомендації щодо ресекторизації, мікроперерв та зміни рівня автоматизації.

Ключові слова: *авіадиспетчер, професійний стрес, система підтримки прийняття рішень, навчання з підкріпленням, пояснювальний штучний інтелект.*

Дата першого надходження статті до видання: 13.03.2026
Дата прийняття статті до друку після рецензування: 10.04.2026
Дата публікації (оприлюднення) статті: 19.05.2026