AI in Flight: Advancing Aviation Safety Through Real-Time Monitoring of Pilots' Neuropsychological States

L. Gicquel, O. Bartheye and L. Fabre
Centre de Recherche de l'École de l'Air (CREA), École de l'Air et de l'Espace, F-13661, Salon-de-Provence, France

Keywords: Artificial Intelligence, Aviation Safety, Neuropsychological Monitoring, Real-Time Systems, EEG, Pilot Cognitive Load, Biometric Sensors, Risk Management.

Abstract: This study proposes an artificial intelligence (AI) system to enhance aviation safety by monitoring pilot neuropsychological states in real-time to address human errors in critical situations. Focusing on light aeronautics as a simplified model, the system aims to monitor and assess the neuropsychological state of the pilots throughout flights, in order to assist the pilots in critical situations, offer solutions, and possibly prevent incidents from happening. Our first approach involves the identification of cognitive factors and electrophysiological correlates that influence pilot performance in extreme conditions. The data is extracted from EEG, ECG, and EOG, combined with camera tracking technologies. This method aims to bridge the current gap between laboratory research and practical application, ensuring that pilots operate safely even in demanding flight conditions.

1 INTRODUCTION

In the domain of aviation safety, a critical and promising area of research is the integration of Artificial Intelligence (AI) autonomous components with real-time monitoring of pilot neuropsychological states to prevent and compensate for human failures in critical situations.

Numerous studies explore these possibility, in autonomous vehicles research as well as aeronautics. However, a significant gap remains between laboratory data, on-terrain recording, and effective use. The present work is an adaptable proof of concept designed using a simple model of light aeronautics, focusing on the paraglider, particularly in its competition (e.g., cross type) or speed variants (i.e., speed-riding or speed-flying, figure 1).

The flexible wing pilot is free from many of the constraints linked to aviation: no engines, less speed, no complex takeoff checklist, although he also faces heightened dangers due to the lack of protection provided by the aircraft, a potential lack of training, and the limitations of his instrumentation. This makes speed-flying an interesting prototypal use-case, concrete and realistic with fewer complexities to take into account. Actually, no system has succeeded in effectively integrating AI agents to proactively manage accident situations and ensure the safety of pilots in extreme conditions. This is the aim of our study, which defines a generic and adaptive concept for exploiting the capabilities of AI to warn and guide pilots throughout the flight and to assist or replace them in the event of a flight failure or incident. To achieve these ends, we detailed a method structured around two key actions: 1) Identify the factors that can influence cognition in an 'extreme' use case and the electrophysiological correlates based on our previous studies (Massé et al., 2022; Melani et al., 2023; Melani et al.,) and 2) Propose an AI algorithm tailored for aeronautics (Lee, 2006) and, a risk diagram and a conceptual analysis of a monitoring system during pilots' activity.

Previous works showed that emotions, cognitive fatigue or mental work load are factors to be taken into account in aviation safety (Dehais et al., 2018; Dönmez and Uslu, 2018; Holtzer et al., 2010; Marcus and Rosekind, 2017; Massé et al., 2022; Melani et al., 2023; Massé et al., 2022). For example, Massé et al. (2022) emphasized the importance of monitoring pilot cognitive workload and cognitive fatigue to prevent inattentional deafness and enhance flight safety. In their experiment, participants had to detect rare sounds in an ecological context of simulated flight under cognitive fatigue. Massé et al. (2022) found that participants performed better to detect alarms under low cognitive load conditions compared with high cognitive load conditions. Results showed that alarm omission and alarm detection can be classified us-

Gicquel, L., Bartheye, O. and Fabre, L.
AI in Flight: Advancing Aviation Safety Through Real-Time Monitoring of Pilots' Neuropsychological States.
DOI: 10.5220/0012977800004562

In *Proceedings of the 2nd International Conference on Cognitive Aircraft Systems (ICCAS 2024)*, pages 80-87
ISBN: 978-989-758-724-5

ing explainable AI and machine learning computation based on time-frequency analysis of brain activity.

Melani et al. (under review) asked participants to verify complex multiplication problems that were either true (e.g., 3×23=69) or false (e.g., 5×98=485). They found that negative emotions modulated some, but not all, arithmetic problem verification mechanisms, with electrophysiological variations observed across different problem types. Interestingly, Melani et al. found that negative emotions influence problem processing from early stages, as evidenced by modulations in early ERP components such as P1 and N2 but also other components such as the P300, the P600 and LPC. In summary, previous studies have shown that cognitive fatigue, mental load and emotions can modify the way in which individuals solve problems.

Thus, under certain conditions, participants may not use the right problem-solving strategies, leading to a drop in performance. These factors not only influence cognitive performance, but may also be associated with changes in electrophysiological components. Consequently, it is essential to take these factors into account and identify them in order to maintain air safety. This is what we do by determining the electrophysical tools that can be used for this particular use case. A comprehensive literature study allowed to outline the physiological characteristics and metrics that can be used for classification and make the connection with the BCI.

The system would record these inputs using a combination of portable and non-invasive methods, ensuring minimal interference with the operator's tasks. Data from EEG, ECG, and PPG are processed using advanced algorithms such as wavelet transforms and machine learning models, including SVM, to accurately classify different psychological and physiological states. Eye and camera tracking technologies provide additional data points for assessing vigilance and cognitive load through fixation metrics, saccades, and pupillary responses.

2 METHODS

The flexible wing pilot is, first and foremost, a "flyer," and the basic principles of flight remain the same. The psychological and physical risks incurred are largely similar, and a psychological or physical failure could lead to a serious incident if it is not taken care of by the pilot himself or by external assistance.

2.1 Determining the Influencing Factors Increasing Risks to the Soft-Wing Pilot in Flight

To explore factors influencing pilots' cognition and decision-making, especially in critical situations, we conducted semi-structured interviews with ten soft wing pilots (min age = 25 years, max age = 60; min flight experience = 50 h, max flight experience = 2000 h). Each interview session lasted between one and two hours, with informal exchange periods.

The pilots' experience ranged from former test pilots to paragliding and speed-flying instructors, as well as leisure pilots with variable experience (between 2 to 15 years). This diverse sample allowed us to examine accident factors and how these factors vary with pilot experience and flight conditions. In this initial phase of our research, we chose not to employ a structured coding strategy, as our main aim was to explore the breadth and depth of pilots' cognitive experiences and decision-making processes in a way that remained open to the rich narratives provided by participants.

This approach allowed us to capture the complexity of pilot experiences without constraining their responses to predefined categories. This narrative method enabled us to identify key themes and patterns that may not have been apparent with a more restrictive coding scheme. Interviews were designed to elicit accounts of routine as well as emergency decision-making processes; pilots recounted memorable flights and incidents as well as flight incident simulation practice detailing their state of mind, action timing and reactions. The questions focused on a three-paneled timeline, before, during and after the flight, to characterize the psychological situations that the pilot undergoes throughout the duration of each phase.

Looking forward, as our research progresses and the initial exploratory findings are further refined, we anticipate utilizing qualitative data analysis software such as NVivo to perform more structured thematic analysis. This will enable us to systematically categorize and analyze the data, providing a means to validate and extend our preliminary insights. Employing NVivo will also facilitate a rigorous examination of the relationships between themes and sub-themes and, help quantify the prevalence of specific experiences and viewpoints, thereby enhancing the robustness and generalizability of our results.

2.2 Electrophysiological Monitoring of the Psychological Factors

To be able to analyze the general situational awareness of the pilot, multiple biometrics sensors can be reviewed. The inputs to the system should include a variety of physiological and behavioral data sources:

- **EEG (Electroencephalography):** For monitoring brain activities, which help assess levels of attention and cognitive load. The mental state is evaluated thanks to EEG micro-states. Brain wave monitoring also allows for specific drowsiness and emotion change detection.
- **ECG (Electrocardiography):** Used to track heart rate variability, which can indicate stress levels.
- **PPG (Photoplethysmography):** For measuring pulse rate variability, providing insights into the pilots' cardiovascular health and stress responses.
- **Eye Tracking / Camera Tracking:** These technologies are critical for evaluating cognitive load and attention by analyzing eye movements, blink rates, and pupil dilation. This helps in understanding how pilots focus and react under different flying conditions.
- **EOG (Electro-oculography) Recording:** This captures eye movement artifacts in addition to EEG data, further enhancing the detection of fatigue and cognitive load.
- **Physiological Indicators:** Measurements of skin conductance and body temperature help in assessing stress levels through changes in the autonomic nervous system.

Most of these wearable sensors are miniaturized. For the EEG recording, we chose to use a dry EEG with 4 electrodes.

2.3 Adaptation of the Design Method to Speedflying

The soft wing paradigm also allows to model quite a variety of flights, as it covers a few types of wings. A very small wing (8-10sqm) will allow for high speed and proximity flying risk modeling, a bigger one (16-19sqm) will allow for longer flights and give more time to the system to propose another outcome than reserve launching. Bigger wings like competitive paragliders can go up to thousands of meters of altitude and allow for very long flights (up to ten hours), which gives space and time to test more complicated scenarios and fatigue or drowsiness onset and subsequent recordings.

3 RESULTS

The initial interview protocol facilitated the identification of various risk factors pertinent to speed-flying, encompassing cognitive load, mental stress, emotional state, drowsiness, and fatigue. Cognitive load denotes the equilibrium between task demands and available time for completion, initially inferred from experimental findings related to problem-solving and learning outcomes, gauged through performance metrics and subjective inquiries. Pilots expressed its impact as "having too much on one's mind," attributing it to personal and professional concerns that diminish attention span and focus capacity. Evaluation of cognitive load may involve EEG or eye-tracking methodologies. Mental stress, partly intertwined with cognitive load, denotes a psychological state arising when individuals perceive demands surpassing coping abilities, instigating psychological and physiological fight-or-flight responses. Participants reported mental stress beyond cognitive load issues, notably in proximity flying, acrobatic maneuvers, or extended flights, often exacerbating during instances of high cognitive load and negative emotions, adversely affecting flight quality and occasionally leading to accidents. Mental stress can precipitate vigilance lapses or misdirected focus. Drowsiness commonly surfaces during extended flights, associated with early waking or insufficient sleep, particularly prevalent in competitive settings. Fatigue emerges post-repeated flights or intense thermal flights, often correlated with larger wing usage, albeit beyond the scope of this preliminary study.

Proposing electrophysiological tools within a BCI system is imperative for assessing and quantifying these factors specific to speed-flying. The BCI system should continuously and unobtrusively monitor pilots' neuropsychological and physiological states during flight, correlating physiological risks with cognitive alterations. Extreme environmental conditions and physiological anomalies during flight significantly impact pilot psychological states. Critical physiological episodes affecting pilots encompass oxygen deprivation, acceleration effects, spatial disorientation, exposure to toxins, and various physiological stresses induced by diverse flying conditions, exacerbated by hypoxia at high altitudes. Additional risks include hypocapnia and hypercapnia affecting cerebral blood flow, extreme gravitational forces leading to G-LOC, and environmental factors like vibrations, temperature fluctuations, and hydration levels. Even minimal instances of these conditions, not reaching pathological levels, could impair pilot decision-making, necessitating consideration.

Figure 1: Speedflying.

3.1 Prototypal Scenario and Risk-Management Strategies

To establish the various scenarios, we conducted extensive research on available accident reports from the French free flight federation (FFVL), videos of flight incidents, and third-party reports, in order to be able to create realistic use-case diagrams. The proposed prototypal scenario is a flight with an obstacle crossing the flight path.

It is to be noted that the soft-wing pilot is freed from many of the constraints associated with aviation: no engines, less speed, no complex take-off checklist. He also faces exacerbated dangers due to the lack of protection provided by the aircraft, a potential lack of training, and the limitations of his instrumentation. The project intends to incorporate risk management strategies from cybersecurity, adapted from methods like EBIOS-RM or CORAS, to address the diverse risks associated with aviation, as well as the system's security itself and the specific risks linked to AI use in safety systems. The CORAS method ((Den Braber et al., 2006), (Stolen et al., 2002)) addresses security-critical systems in general, but places particular emphasis on IT security. An IT system for CORAS is not just technology, but rather a medium for communication and interaction between different groups of stakeholders involved in a risk assessment; what matters is the human interacting with the technology and all relevant aspects of the surrounding organization and society (Table 1).

Risks are events that harm assets when they occur. However, often some risks are accepted, either because of shortage of resources or because of conflicting concerns.

Scales for the likelihood of which incidents occur are defined (certain, likely, possible, unlikely, rare) including the impact or consequence they have on the assets. Assets are ranked according to their importance to distinguish risks that can be accepted from those that cannot.

During the requirements phase, designers and pilots establish the need for a system and document its purpose. Security planning should begin at the requirements phase and consist of activities to establish security requirements and assess security needs. The security risk assessment helps us to define the functional characteristics of the system requiring a security need. During this stage, teams also determine the need for threat modeling, reviews, and security design. At the requirements stage, availability and integrity of system services are taken into account as well as privacy and data sensitivity measures. Requirement diagrams (see figure 2) define the technical (hardware and software) requirements of a system capable of real-time monitoring of pilots' neurological and physiological state. BPMN diagrams (figure 3) represent the various flight scenarios where the system should monitor the pilot biometrics, and when it should intervene or interact with the pilot.

4 DISCUSSION

The present concept is based on multiple data inputs. Considering there is a need for simplification of the data stream to allow for real-time processing, we intend to implement camera eye tracking and EOG analysis to classify the cognitive load of the pilot on first intention, as EEG data could be primarily used for the detection of drowsiness, fatigue or stress.

The cognitive load analysis involves eye tracking

Table 1: IT system.

Initiator	**Incident** Threat scenario, Unwanted incident	**Cause** Vulnerability	**Risk** Risk	**Treatment** Treatment scenario
Threat (accidental) (Pilot)	Obstacle not seen; can cause a collision	Drowsiness / Deficient mental state	Collision	Allow AI device to take control
Threat (non-human) (AI)	Emotional state not detected (false negative); cannot detect deficient mental state	Inefficient machine learning model or incomplete dataset	Unrecovered human failure	Improve model and dataset

to monitor where the pilot directs his gaze, and correlation to EEG data to analyse his information processing. Vigilance and attention can thus be assessed by a combination of EEG and eye tracking (EOG recording and/or camera tracking). Consequently, our concept aims to quantify precisely specific eye movement and qualify an increase in blinking and slower eye movements as signs of fatigue. Constant eye focus and a lack of blinking would, conversely, indicate high levels of vigilance and attention.

While it is possible to quantify blinking using cameras only, the modularity of this system could allow to maintain the quality of the assistance when the use of cameras is impossible. To account for physiological changes and to support this data for neurophysiological state assessment, we intend to complement this setting with GSR (skin conductance and temperature), ECG (heart rate and variability), and possibly plethysmography data, which measures changes in volume in different parts of the body.

4.1 Computational Data Analysis

On the computational level, according to Al-Shargie and al.(Al-shargie et al., 2016), as EEG signals are non-stationary, wavelet transform can provide clear features that are strongly correlated with levels of mental stress. Since this study, several other teams have successfully used Support Vector Machine (SVM) algorithms and other machine learning methods to characterize stress, generally achieving detection levels around 95%, especially when combining multiple models. Another possibility could be to use non-supervised methods.

An interesting work on this subject describes a potential contribution of artificial intelligence and deep learning (Lee et al., 2023) to the flight environment. The study presents an autonomous system for EEG-based multiple abnormal mental states classification using hybrid deep neural networks. Various specific abnormal mental states (namely, low fatigue, high fatigue, low workload, high workload, low distraction, and high distraction) are classified by applying the deep learning method. The accuracy is, as customary with deep learning networks, quite high, but a second line of classifying, based on traditional algorithmic rules, could be applied to make the output of the deep learning model robust and secure in critical cases.

4.2 System Outputs and Performance

The first model we have chosen for the sake of simplicity is the mid-range wing, around 16 sqm. Even if in some cases (thermal flights) the aircraft could theoretically have the altitude and remaining flight time for the system to propose various solutions (calling an operator, taking some time to wake the pilot), speed-flying accidents usually go fast.

As such, the preferred output of the system here would be, as soon as it is detected that the pilot is unresponsive, to fire a pyrotechnic parachute system (which have very recently been introduced to the soft wing world). More complicated outputs will be assessed by modelling incidents with bigger aircrafts and/or longer flights. We could then establish a connection between an external operator or a discussion between a specifically trained large language model (LLM) to guide the pilot to landing or keep him aware enough to solve the situation.

The outputs of the system should provide a comprehensive assessment of the operator's readiness and condition. These outputs include:

- **State Detection:** Identifying levels of cognitive load, stress, and attention states, which allows for timely interventions to prevent accidents or errors.
- **Risk Prediction:** Forecasting potential physio-psychological risks such as extreme fatigue or high stress, which could impair performance.
- **Adaptive Feedback:** Offering real-time, adaptive feedback to the operator or automated systems on board, which can adjust the workload or provide alerts to mitigate risks.

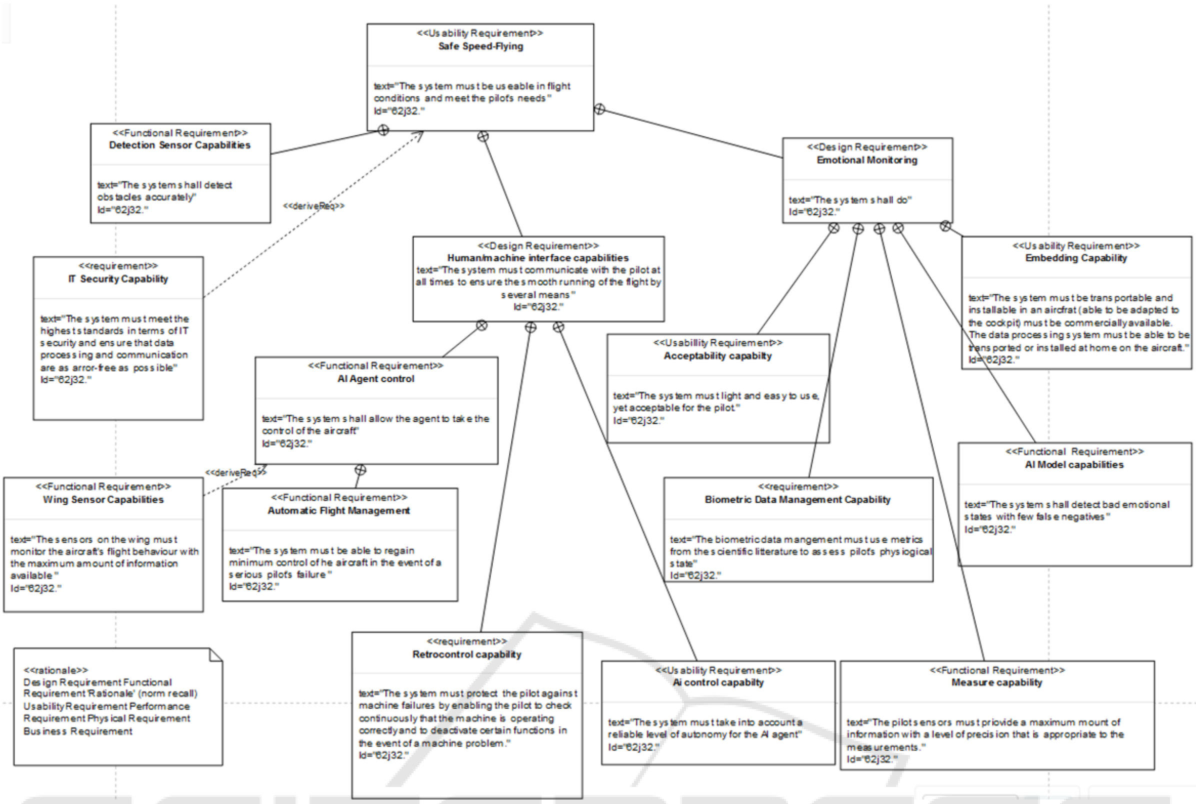

Figure 2: Requirement diagram of the speedflying system.

In the speedflying use case, often, the best action of the system would be to prevent the flight scenario from occurring entirely. A public report for a deadly accident in 2022 from FFVL presented this as a conclusion. One of the influencing factors in the incident was that "The pilot had not had a good night's sleep and was reluctant to go flying." These kind of situations should be taken into account by a general health assessment by the biometrics sensors as well as some routine questions before flight, which could be formulated in a natural speech style by a trained LLM.

In case of a strongly dangerous situation, the system should aim to prevent the pilot from flying by informing him and, possibly, calling a person that would be pre-designated as a safeguard, to try and discourage the pilot from flying on this particular instance. This system thus acts as a critical tool in enhancing the operational safety and efficiency of pilots and other aerospace operators by providing a real-time, integrated view of their psycho-physiological status and adapting to their immediate needs.

The risk-management method adapted here to a specific prototypal use-case, speedflying, is extensible to any type of aeronautical use-case, by adding new actors, taking into account a larger array of influencing factors as well as the specifics of the considered aircraft and types of flight. Specific physiological factors linked to extreme altitude, life support systems or pressurization of the cabin could be added in the assessment. Fatigue and drowsiness, which are generally linked to longer flights, could also be monitored with the same sensor setup. For example, drowsiness can be detected with a high accuracy by EEG (Balandong et al., 2018). The recording area of preference would, in this case, be the occipital cortex; we could dedicate a limited number of electrodes for this task.

In conclusion, this BCI system designed from the knowledge collected is of real interest in improving the safety of paragliding flights and is intended to be generalized to other forms of flight (Deng and others, 2020). The groundwork is set for the proposal of a similar device that would function on a more complicated use-case, military aviation. Future efforts will concentrate on developing and training AI algorithms to process and interpret complex data, as well as gaining pilot acceptance of this kind of products. The creation of a first simplified prototype based on this research would be possible to this effect.

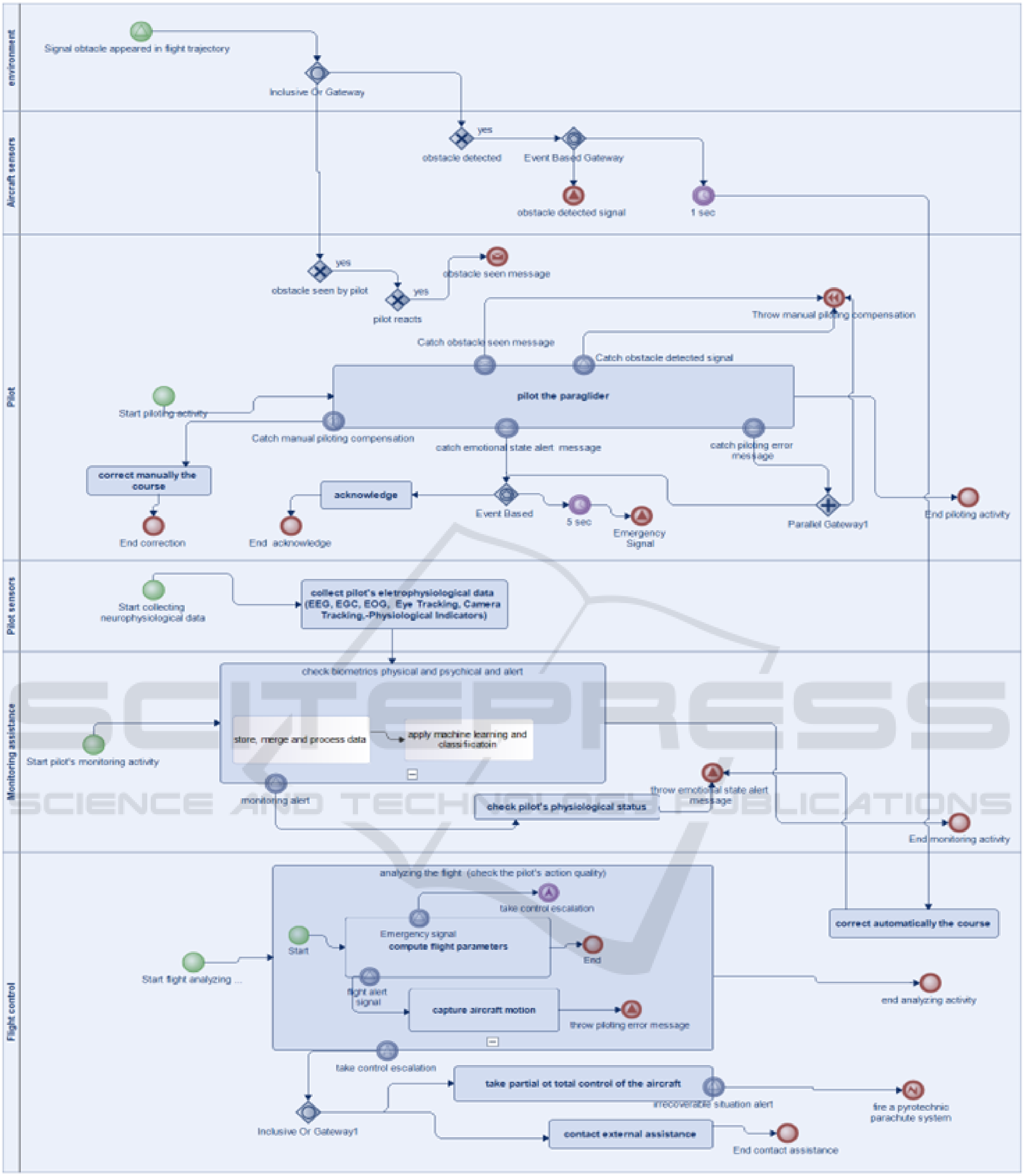

Figure 3: BPMN diagram of a flight with an obstacle scenario.

REFERENCES

Al-shargie, F. M., Tang, T. B., Badruddin, N., and Kiguchi, M. (2016). Mental Stress Quantification Using EEG Signals. In Ibrahim, F., Usman, J., Mohktar, M. S., and Ahmad (Éds, M. Y., editors, *International Conference for Innovation in Biomedical Engineering and Life Sciences*, pages 15–19. Springer.

Balandong, R. P., Ahmad, R. F., N., M. S. M., and Malik, A. S. (2018). A Review on EEG-Based Automatic Sleepiness Detection Systems for Driver. *IEEE Access*, 6:22908–22919.

Dehais, F., Dupres, A., Di Flumeri, G., Verdiere, K., Borghini, G., Babiloni, F., and Roy, R. (2018). Monitoring pilot's cognitive fatigue with engagement features in simulated and actual flight conditions using an hybrid fNIRS-EEG passive BCI. In *2018 IEEE international conference on systems, man, and cybernetics (SMC)*, pages 544–549. IEEE.

Den Braber, F., Brændeland, G., Dahl, H. E., Engan, I., Hogganvik, I., Lund, M. S., others, and Vraalsen, F. (2006). The CORAS model-based method for security risk analysis. *SINTEF, Oslo*, 12:15–32.

Deng, P.-Y. and others (2020). Detecting fatigue status of pilots based on deep learning network using EEG signals. *IEEE Transactions on Cognitive and Developmental Systems*, 13(3):575–585.

Dönmez, K. and Uslu, S. (2018). The relationship between flight operations and organizations in aircraft accidents; the application of the human factor analysis and classification system. *Anadolu University Journal of Science and Technology A - Applied Sciences and Engineering*, pages 1–1.

Holtzer, R., Shuman, M., Mahoney, J. R., Lipton, R., and Verghese, J. (2010). Cognitive Fatigue Defined in the Context of Attention Networks. *Aging, Neuropsychology, and Cognition*, 18(1):108–128.

Lee, D. H., Jeong, J. H., Yu, B. W., Kam, T. E., and Lee, S. W. (2023). *Autonomous system for EEG-based multiple abnormal mental states classification using hybrid deep neural networks under flight environment*. Systems, IEEE Transactions on Systems, Man, and Cybernetics.

Lee, W.-K. (2006). Risk assessment modeling in aviation safety management. *Journal of Air Transport Management*, 12(5):267–73.

Marcus, J. H. and Rosekind, M. R. (2017). Fatigue in transportation: NTSB investigations and safety recommendations. *Injury prevention*, 23(4):232–238. Publisher: BMJ Publishing Group Ltd.

Massé, E., Bartheye, O., and Fabre, L. (2022). Classification of electrophysiological signatures with explainable artificial intelligence : The case of alarm detection in flight simulator. *Frontiers in Neuroinformatics*, 16.

Melani, P., Fabre, L., and Lemaire, P. *How negative emotions influence arithmetic performance : a comparison of young and older adults*.

Melani, P., Fabre, L., and Lemaire, P. (2023). How negative emotions influence arithmetic performance : a comparison of young and older adults. *Current Psychology*.

Stolen, K., Den Braber, F., Dimitrakos, T., Fredriksen, R., Gran, B. A., Houmb, S. H., others, and Aagedal, J. O. (2002). Model-based risk assessment–the CORAS approach. In *iTrust Workshop*.