

Article

Not peer-reviewed version

Quantitative Assessment of Drone Pilot Performance

[Daniela Doroftei](#)*, [Geert De Cubber](#), Salvatore Lo Bue, Hans De Smet

Posted Date: 24 July 2024

doi: 10.20944/preprints202407.1957.v1

Keywords: Performance assessment; Pilot training; Human performance modelling; Quantitative validation; Standardised tests; Drone simulator; AI copilot



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Quantitative Assessment of Drone Pilot Performance

Daniela Doroftei ^{1,*}, Geert De Cubber ¹, Salvatore Lo Bue ² and Hans De Smet ³

¹ Royal Military Academy of Belgium, Robotics & Autonomous Systems unit; daniela.doroftei@mil.be (D.D.); geert.de.cubber@mil.be (G.D.C.)

² Royal Military Academy of Belgium, Department of Life Sciences; salvatore.lobue@mil.be

³ Royal Military Academy of Belgium, Department of Economy, Management and Leadership; hans.de.smet@mil.be

* Correspondence: daniela.doroftei@mil.be; Tel.: +32-2-4414106

[†] Current address: Avenue de la Renaissance 30, Brussels 1000, Belgium.

[‡] This article is based on the research work of the first author. The second, third and fourth author supervised this work, gave guidance and reviewed the article.

Abstract: This paper introduces a quantitative methodology for assessing drone pilot performance, aiming to reduce drone-related incidents by understanding the human factors influencing performance. The challenge lies in balancing evaluations in operationally relevant environments with those in a standardized test environment for statistical relevance. The proposed methodology employs a novel virtual test environment that records not only basic flight metrics but also complex mission performance metrics, such as the video quality from a target. A group of Belgian Defence drone pilots were trained using this simulator system, yielding several practical results. These include a human performance model linking human factors to pilot performance, an AI co-pilot providing real-time flight performance guidance, a tool for generating optimal flight trajectories, a mission planning tool for ideal pilot assignment, and a method for iterative training improvement based on quantitative input. The training results with real pilots demonstrate the methodology's effectiveness in evaluating pilot performance for complex military missions, suggesting its potential as a valuable addition to new pilot training programs.

Keywords: performance assessment; pilot training; human performance modelling; quantitative validation; standardised tests; drone simulator; AI copilot

1. Introduction

1.1. Need and Problem Statement

As the number of drone operations rises each year, so does the risk of incidents with these unmanned systems [1,2]. Research shows that the majority of these incidents can be attributed to human error [3]. From a human factors perspective, improving human-machine interaction can be performed through improving the training, the equipment, the task definition, the environment, the selection of the personnel or the organisation of the work. In this paper, we will focus on investigating the root causes of these errors and integrating these findings into training procedures. While acknowledging the various human factors that can improve human-machine interactions [4], we chose to focus on enhancing training, because, unlike the improvement of systems or of tasks on these systems, training is an area over which the armed forces have direct control. However, assessing the human factors contributing to drone incidents presents significant challenges. On one hand, a realistic operational environment is necessary to study human behaviour under authentic conditions. On the other hand, a standardized environment is essential for conducting repeatable experiments that ensure statistical relevance of the results.

Generally, two approaches are distinguished towards quantitative assessment of drone pilot performance [5]: detailed incident report analysis and simulation environments. However, for small drone operations, incident reporting is often inadequate, lacking sufficient data for meaningful analysis. Simulation environments on the other hand do offer controlled settings to observe pilot behaviour, which is a common practice in manned aviation where extensive simulator training precedes real flights [6]. However, for small drones, achieving realistic sensory feedback is challenging, as the

current simulators are often limited to simplistic scenarios, failing to provide high-quality feedback to trainees or mentors [7].

To address these issues, we propose a drone operator performance assessment tool that utilizes a realistic environment and operational conditions to measure operator performance qualitatively and quantitatively. The ultimate goal is to use these metrics to optimize training curricula, ensuring maximum safety during real operations.

Realism is crucial in simulation systems to achieve desired results. Our proposed framework uses a highly realistic environment, incorporating operational conditions like wind and weather effects. A key aspect of qualification assessment is defining test methods and scenarios. Current drone training simulators often cover only simple operations, while complex scenarios reveal critical human factors. For this reason, pilots in challenging conditions, such as military or emergency services, find simplistic scenarios irrelevant. Therefore, this paper introduces a standard test methodology specifically for assessing drone operators in the security sector.

Our performance assessment tool includes a simulator for military and other drone pilots to perform missions in realistic virtual environments. Developed for this research, the simulator tracks pilot performance parameters, offering a highly realistic setting for complex operations. Built on the Microsoft AirSim engine [8] and the Unreal Engine [9], the simulator is customizable, enabling the incorporation of standard test scenarios and multiple drone models. It measures over 65 flight parameters and assesses video footage quality, crucial for evaluating operator performance. The tool also suggests optimal trajectories for target observation, assisting operators during missions.

Post-flight, an overall performance score is calculated based on mission objectives. After subjecting a range of pilots (from beginners to experts) to the simulator environment, the obtained data was used to:

1. Train an AI-based classifier to recognize 'good' and 'bad' flight behaviour in real-time, aiding in the development of a virtual AI co-pilot for immediate feedback;
2. Develop a mission planning tool assigning the best pilot for specific missions based on performance scores and stress models;
3. Identify and rank human factors impacting flight performance, linking stress factors to performance;
4. Update training procedures to account for human factors, improving pilot awareness and performance.

In summary, this paper introduces the following novel contributions:

- A drone simulator tracking operator performance under realistic conditions in standardized environments (discussed in Section 3).;
- A method for quantitative evaluation of drone-based video quality (discussed in Section 3.3);
- A methodology for modelling human performance in drone operations, relating human factors to operator performance (discussed in Section 4);
- An AI co-pilot offering real-time flight performance guidance (discussed in Section 5.1);
- A flight assistant tool for generating optimal flight trajectories (discussed in Section 5.2);
- A mission planning tool for optimal pilot assignment (discussed in Section 5.3);
- An iterative training improvement methodology based on quantitative input (discussed in Section 5.4).

This paper discusses these contributions in detail, starting with a review of related work in the remainder of Section 1, followed by an introduction to the general framework in Section 2. Sections 3 and 4 concentrate on the core contributions introduced by this work: on the one hand the novel simulator tool and on the other hand the pilot performance modelling methodology. By training a series of pilots (following the methodology described in section 4) in this simulator environment (described in section 3), these innovations are combined, which allows for the development of a number of applications that are described in section 5.

1.2. Related Work in the Domain of Drone Simulation

Drone simulators have become integral tools for both recreational users and professional operators, offering a risk-free environment for training and experimentation. A detailed review of the performances of these systems has been performed by Mairaj et al. [10]. For the sake of brevity, this section focuses on the most prominent simulators that are currently available, including the DJI Drone Simulator [11], the DRL Sim Drone Racing Simulator [12], the Zephyr Drone Simulator [13], the droneSim Pro Drone Flight Simulator [14], and the RealFlight RF9.5 Drone Simulator [15]. This section provides an overview of these simulators, discussing their features, advantages, and weaknesses, and comparing their utility for different user needs.

The DJI Drone Simulator [11] is primarily designed for training pilots to operate DJI drones, offering both free and paid versions, with the latter providing more advanced features and scenarios. One of the main advantages of the DJI Drone Simulator is its realistic flight physics, which closely replicate the characteristics of actual DJI drones. The simulator also offers a variety of environments, from urban landscapes to open fields, and includes comprehensive training modules and scenarios, which are particularly useful for new pilots. However, the simulator is limited to DJI drones, reducing its utility for users operating other brands. Additionally, the advanced features and realistic environments require a paid subscription, which may not be affordable for all users.

The DRL Sim Drone Racing Simulator [12], developed by the Drone Racing League (DRL), focuses on high-speed drone racing and aims to provide an authentic racing experience. The simulator's physics engine provides a realistic racing experience that closely mimics actual drone racing conditions. Users can compete against others online, adding a social and competitive dimension, and a wide range of race tracks and customizable courses are available. The DRL Sim includes tools for fine-tuning drones and improving racing skills. However, its high specialization in racing makes it less useful for those interested in other aspects of drone operation, and the focus on high-speed manoeuvres and precision can be challenging for beginners.

Zephyr Drone Simulator [13] is designed primarily for educational purposes, providing a platform for both recreational and professional training. It emphasizes training and education, featuring structured lessons and comprehensive feedback. Zephyr supports a variety of drone models, making it versatile for different types of training, and allows instructors to create custom training scenarios tailored to specific needs. Despite these advantages, Zephyr's graphical quality and realism may not be as high as some other simulators, and it may lack some advanced features sought by experienced pilots.

droneSim Pro Drone Flight Simulator [14] offers a realistic flight experience aimed at both hobbyists and professional pilots. It focuses on providing realistic flight physics and dynamics, and its straightforward interface makes it accessible to beginners. Additionally, it is relatively affordable compared to some other high-end simulators. However, droneSim Pro's variety of training environments and scenarios is not as extensive as some competitors, and its graphics may not be as advanced, impacting the overall immersive experience.

RealFlight RF9.5 Drone Simulator [15] is a comprehensive flight simulator that includes a variety of aircraft, including drones. It supports a wide range of aircraft types, making it useful for users interested in more than just drones. The simulator offers high-quality graphics and a realistic flying experience, with extensive training tools and scenarios catering to all skill levels. The strong community support and regular updates enhance the user experience. However, the broad range of features and aircraft can be overwhelming for beginners, and the simulator is on the pricier side, which might be a deterrent for some users.

While each of these simulators has unique strengths, they also share a common limitation: the inability to record pilot performance and flight characteristics during flights. This deficiency impedes their utility for the quantitative assessment of drone pilot performance, restricting their potential for comprehensive training and evaluation.

1.3. Related Work in the Domain of Quantitative Evaluation of Drone Pilot Performance

The necessity for a quantitative assessment of drone pilot performance, initially observed in military contexts by Cotting [16] and Holmberg [17], has led to the application of various evaluation techniques [18,19]. These include traditional analytical performance evaluations [17,20] and the increasingly popular mission-task-elements (MTE) paradigm, favoured for its flexibility and universality [21–23].

Efforts have been made to characterize different types of Unmanned Aerial Systems (UAS) and establish databases for performance criteria of existing systems [16,24]. A methodology has been proposed for the systematic development and refinement of MTEs, which involves the use of simulated vehicle performance checked against flight test data to refine the model until it achieves sufficient accuracy [19,22].

Building on this, Herrington et al. explored in [25] the human factors associated with training pilots to meet the demands of commercial drone operation. Ververs and Wickens [26] conducted experiments to study the impact of multiple interface variables on pilot performance in a cruise flight environment. They used statistical analysis methods such as mean absolute error, root mean square error (RMSE), and response time delay. The experiments involved pilots flying a flight segment while observing changes in heading, airspeed, and altitude, and pressing a button on the joystick when any randomly appearing external event was detected. The deviations from the flight path and response time to detect sudden events were measured using the aforementioned statistical methods, and the experiments were repeated for different display locations, clutter, and image contrast ratios. The data suggested that attention was modulated between tasks (flight control and detection), and between display areas (head-up and head-down).

Smith and Caldwell [27] employed RMSE to investigate pilot fatigue in extensive experiments involving many turns and climbs before performing an instrument landing. The study suggested the use of objective metrics such as RMSE over predefined flight patterns for measuring pilot performance. Hanson et al. [28] evaluated the usability of different adaptive control techniques by using pilot handling quality ratings for an air-to-air tracking manoeuvre. The ratings were computed for pitch stick deflections from two pilots while using different adaptive controllers and with simulated aircraft failures.

Field and Giese [29] further applied RMSE and power spectral density (PSD) of pitch control column activity during handling quality research investigations for a lateral offset approach and landing task. The focus of this study was to evaluate the changes in the handling qualities for the same aircraft configuration in three different simulators. The authors conclude that primarily the PSD metric is useful to qualitatively and quantitatively assess the pilot's control activity input power.

Other researchers, such as Zahed et al. [30], have focused on Cross-track command and path error as potential time-domain metrics to quantify pilot and quadcopter performance to establish training and certification for pilots and aircraft. Archana Hebbar and Pashilkar [31] examined the application of different statistical and empirical analysis methods to quantify pilot performance. They executed a realistic approach and landing flight scenario using the reconfigurable flight simulator at the Indian National Aerospace Laboratories and applied both subjective and quantitative measures to the pilot performance data. Results indicate that analysing pilot's control strategy together with his/her deviations from predetermined flight profile provides a means to quantify pilot performance.

In summary, while a variety of statistical measures are available in the literature, specifying a common metric to measure performance is challenging due to context specificity. A range of metrics has been developed to assess different performance dimensions, but these are often task-specific and rely on measuring the deviation from a predefined ideal trajectory. This approach is less useful when the pilot is tasked with complex missions, such as searching for missing persons or intruders in an area, where no ideal trajectory can be defined.

An example of an approach that does take into consideration more complex mission profiles is the methodology proposed by the National Institute of Standards and Technology (NIST) [32].

NIST developed quantitative assessment methods for drones and their operators, particularly in the context of urban search and rescue operations. This research from NIST includes standardized testing methods and infrastructure to evaluate the performance of drone operators and enables users to generate statistically significant data on aspects such as airworthiness, manoeuvrability, and payload functionality. Although these methods are valuable, they are heavily focused on urban search and rescue operations and may not be applicable to all types of security missions.

Consequently, there remains a significant need for reliable quantitative evaluation methodologies for drone pilot performance in various complex mission profiles.

1.4. Related Work in the Domain of Video Quality Analysis

Assessing video quality in drone operations is crucial for ensuring accurate navigation, target identification, and situational awareness, which are essential for mission success and safety. High-quality video feeds enable operators to make informed decisions and respond promptly to dynamic environments. Therefore, an essential component of the proposed simulator system is a tool to quantitatively assess the video quality.

Video quality analysis methodologies can be broadly classified into two categories: subjective and objective methods.

Subjective video quality analysis methodologies [33] evaluate video quality as perceived by humans. These methods require video sequences to be shown to groups of viewers, whose subjective opinions are recorded and averaged into a Mean Opinion Score (MOS) to determine the quality of the video sequence. Although subjective video analysis methods provide excellent results, they are extremely labour-intensive and thus challenging to implement in practical contexts. Consequently, more objective—and therefore more automated—methods have been developed.

Objective video quality analysis methods are classified by the International Telecommunications Union (ITU) based on the input data used by the algorithms [34], as shown on Figure 1.

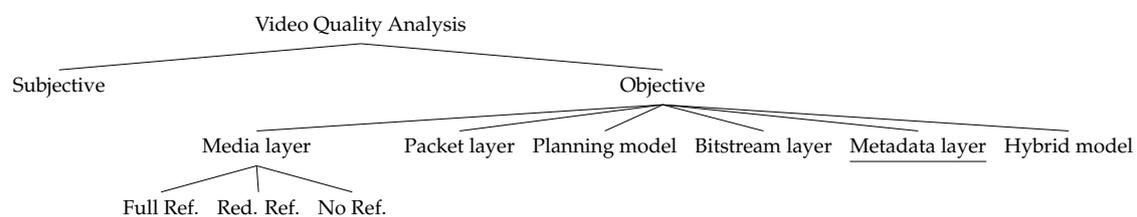


Figure 1. Taxonomy of video quality analysis methodologies.

Media layer models, for instance, directly utilize the video signal to define a quality measure. These methods do not require prior information about the system under test and are often used to compare different compression methodologies. Depending on the type of source data to which the processed video is compared, three sub-methods can be identified:

1. Full-Reference Methods: These extract data from high-quality, non-degraded source signals and are often derivatives of PSNR [35], commonly used for video codec evaluation;
2. Reduced-Reference Methods: These extract data from a side channel containing signal parameter data;
3. No-Reference Methods: These evaluate video quality without using any source information.

Parametric packet-layer models estimate video quality using only information from the signal packet header.

Parametric planning models use quality planning parameters for networks to estimate video quality.

Bitstream layer models combine encoded bitstream information with packet-layer information to estimate video quality.

Hybrid models use a combination of these methodologies to estimate video quality.

Traditional methodologies for video quality analysis are based on the assumption that there is a perfect input signal, which is then degraded due to encoding, network transmission, decoding, and display constraints. However, in our application, this assumption does not hold. Our primary interest lies in whether specific subjects are adequately perceived within the video material, necessitating a content-based analysis.

Content-based video analysis for drones has been explored by Hulens and Goedemé [36], who presented an autonomous drone that automatically adjusts its position to keep a subject (e.g., an interviewee) within view under certain cinematographic constraints. While this approach is useful for specific applications where subjects are always human faces, our research aims to maximize information gain about generic subjects.

The full-frame analysis proposed by Hulens and Goedemé [36] imposes significant constraints on system processing requirements. These constraints can be avoided by using only metadata for analysis. Drones typically have accurate GPS sensors onboard, allowing for the geolocation of all image and video data produced. This enables the introduction of a new category of models based on metadata layer processing [37].

The new metadata-layer model, however, overlooks crucial aspects of the traditional video quality analysis paradigm, such as errors that may occur in the encoding-transmission-decoding-display pipeline. However, for a fact and accurate assessment of the quality of the content of the produced video by a drone pilot (while ignoring eventual transmission losses), the metadata approach is highly appealing.

1.5. Related Work in the Domain of Human Performance Modelling for Drone Operations

The study of human performance in drone operations has evolved significantly over the years. Early research by the US Air Force in 2006 focused on large, remote-controlled military drones, with performance models primarily analysing operator workload to optimize crew composition [38]. However, these models may not be applicable to smaller drone systems with limited crews.

In response, Bertuccelli et al. [39] introduced a new formulation for a single operator conducting a search mission with multiple drones in a time-limited environment. This concept was further developed by Wu et al. [40], who proposed a multi-operator multi-drone model.

The main conclusion of the research cited above was that cognitive workload and stress factors significantly affect drone pilots performance. High cognitive workload and stress can impair a pilot's decision-making and reaction times. Therefore, tools such as NASA's Task Load Index (NASA-TLX) [41] were introduced in the domain of human performance modelling for drone operations to assess perceived workload during drone operations.

A more recent innovation is the integration of autonomous systems with human operators, which brings with it a whole new series of open research questions related to human performance modelling for drone operations [42]. Effective human-autonomy teaming can reduce cognitive workload and improve performance [43]. However, challenges remain in designing interfaces and interaction protocols that maximize the benefits of automation while minimizing new types of errors.

AI and machine learning techniques can also be applied to model and predict human performance in drone operations. AI-based systems can analyse large datasets from simulator and real-world operations to identify patterns of behaviour associated with high and low performance [44]. Machine learning models can predict operator errors and provide real-time feedback or intervention.

Advancements in wearable technology have enabled the real-time monitoring of physiological indicators such as heart rate variability, galvanic skin response, and brainwave activity [45]. These indicators can be used to assess an operator's stress and fatigue levels [46]. Biofeedback mechanisms can then be employed to help operators manage their stress and maintain optimal performance levels during drone operations.

Developing standardized test scenarios and performance metrics is crucial for assessing and improving human performance in drone operations [47]. Standardization allows for consistent evaluation across different operators and operational contexts. Metrics such as mission completion time, error rates, and quality of data collected are used to measure performance. These metrics help in identifying areas where training and operational procedures can be improved.

While significant progress has been made in understanding and improving human factors in drone operations, challenges remain in achieving realistic simulation environments, developing effective human-autonomy teams, and creating comprehensive training programs. Specifically for military operations, critics argue that existing methods over-rely on classical attention and fatigue modelling, while ignoring essential elements for security sector operations, such as mission stress, enemy countermeasures, and varying operator skills. This highlights the need for further research and development in this field.

1.6. Related Work in the Domain of AI Assistance for Drone Operations

The proposed work includes the development of an advanced pilot assistance system (APAS) for drone operations, based on the assessment of the performance of human operators in a virtual flight simulator environment.

Existing APAS for drone operations are already equipped with several advanced features [48]. Many commercial drones possess the capability to return home automatically and detect and avoid obstacles. These systems utilize multiple sensors, leverage artificial intelligence (AI) for data analysis, and incorporate advanced control algorithms to assist pilots in various aspects of drone operations [49].

One of the primary areas of development in pilot assistance systems is autonomous navigation and obstacle avoidance. These systems use a combination of sensors, such as LiDAR, cameras, and ultrasonic sensors, to detect and avoid obstacles in real-time [50]. Advanced algorithms, including Simultaneous Localization and Mapping (SLAM) and path planning algorithms, enable drones to navigate complex environments autonomously [51]. SLAM algorithms allow drones to create and update maps of unknown environments while simultaneously tracking their own location. This technology is crucial for indoor and GPS-denied environments. Path planning algorithms, such as Rapidly-exploring Random Trees and A*, enable efficient path planning by finding the shortest and safest routes around obstacles [52].

AI-based flight control systems enhance the stability and control of drones, especially in challenging conditions. Machine learning algorithms can predict and compensate for environmental disturbances such as wind gusts, ensuring smoother flights [53]. Reinforcement learning algorithms have shown promise in optimizing flight control by learning from interactions with the environment [54]. These systems continuously improve their performance through trial and error. Neural networks are used to model complex dynamics and provide adaptive control strategies, making drones more resilient to varying flight conditions.

Computer vision technologies enable drones to interpret and analyse visual data, providing critical assistance in tasks such as target tracking, landing, and navigation. Advanced algorithms, such as YOLO (You Only Look Once) and Faster R-CNN (Region Convolutional Neural Network), allow drones to detect and track objects in real-time, enhancing their capability to follow moving targets or avoid dynamic obstacles [55]. Visual odometry techniques use camera data to estimate the drone's position and movement, which is particularly useful for GPS-denied environments [56].

Effective Human-Machine Interface (HMI) designs are crucial for providing intuitive control and feedback to drone pilots. Advances in HMI focus on improving situational awareness and reducing pilot workload [57]. Augmented Reality (AR) interfaces overlay critical flight information onto the pilot's field of view, enhancing situational awareness without the need to look away from the drone [58]. Haptic feedback systems provide tactile sensations to the pilot, indicating obstacles, flight status, or other critical alerts, thereby improving control precision [59].

Autonomous landing systems are designed to assist drones in safely landing, particularly in challenging environments or emergencies [60]. Systems equipped with computer vision and GPS technologies enable drones to land accurately on predefined markers or docking stations. Emergency landing algorithms identify safe landing spots in real-time during emergencies, such as power failures or loss of communication [61].

While most of these APAS solutions focus on automated guidance and exteroceptive sensing for collision avoidance, the aspect of human-drone interaction during flight operations is less explored. However, recent developments are emerging to provide real-time feedback to pilots about their flight performance. These systems use data analytics and AI to evaluate various performance metrics and offer actionable insights [62].

Performance monitoring systems track parameters such as flight stability, adherence to planned routes, and response to environmental conditions. Real-time analytics allow pilots to receive immediate feedback on their actions, helping them to adjust their behaviour during flight [44]. For instance, if a pilot exhibits flight patterns indicative of fatigue or stress, the system can provide warnings and suggest corrective actions.

An example is the work performed within the DLR MOSES (More operational flight safety by enhancement of situation awareness) project [63], where different approaches to measuring and improving situation awareness were investigated, particularly in information gathering during approach and taxiing. The study analysed eye movements of forty pilot students and nine experienced pilots and explored the influence of an additional taxi guidance system on eye movements. While these results are significant, they primarily focus on ground operations rather than flight operations.

AI-driven co-pilot systems have been developed for manned aviation to assist pilots by offering real-time guidance and support. These systems analyse flight data in real-time, comparing it against optimal flight models, and provide suggestions or corrections to the pilot [64]. Such systems enhance safety by reducing the likelihood of human error and improving overall flight performance.

2. Overview of the Proposed Evaluation Framework and Situation in Comparison to the State of the Art

As stated in the previous sections, there has been some previous work in the domain of quantitative assessment of drone pilot performance. However, an all-encompassing solution that provides an integrated approach towards human performance modelling in a statistically relevant and standardised training environment, is lacking.

The concept of the approach presented in this paper is sketched in Figure 2 and is based on the incorporation of a novel drone simulator tool that is used for human performance modelling. The ultimate goal of this architecture is to provide quantitative performance feedback as input to a pilot training, such that corrective measures can be taken if required and that the training curricula can be improved.

As discussed in section 1.2, current drone simulation tools available on the market do not enable the recording of pilot performance and flight characteristics during the flight. We propose to tackle this problem, by developing a simulator tool that is built on the Microsoft AirSim engine [8], which is an Open source simulator for autonomous vehicles built on the Unreal Engine [9]. Building on top of this open source tool, we integrate complex standardised test scenarios that cater to the needs of military security professionals. In our use case, we consider specifically the needs of the Belgian Special Forces Group, and for that purpose, we focus on a scenario of intruder detection and identification, but the framework is totally open to deal with different use case scenarios.

An important aspect of such a use case - and in many others, such as any scenario involving photography or mapping - is evaluating the quality of the video processed by the drone pilot. Within the proposed simulator, we have integrated a unique content-based video-analysis tool based, based on metadata analysis. This video quality analysis approach was first introduced in our previous work [37]

as an isolated video quality assessment methodology, and in this paper we show how it is integrated in a full pilot performance assessment pipeline.

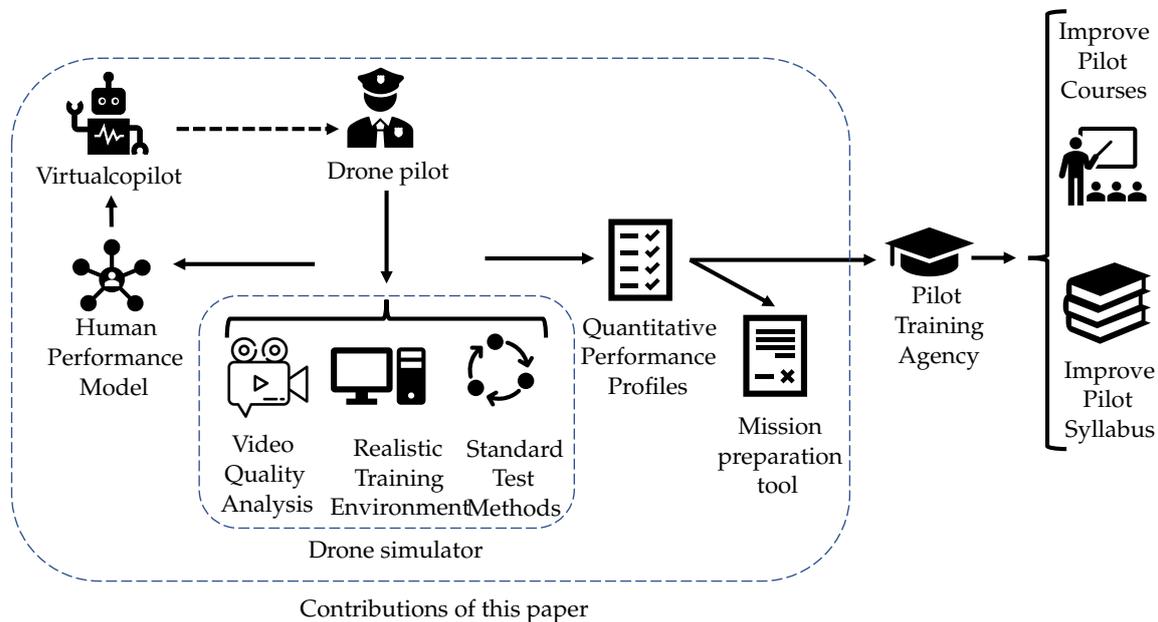


Figure 2. Conceptual overview of the main contributions of this paper.

In the domain of human performance modelling, the presented approach builds on existing work [65,66] on qualitative performance modelling (using standardised questionnaires) and quantitative performance modelling (using training in highly realistic and standardised simulation environments). Furthermore, this paper introduces and discusses the performance modelling results from extensive trials with civilian and military test pilots working for Belgian Defence. These results led to the development of a drone mission planning tool that enables to select the optimal pilot and drone for a given mission in function of the performance results recorded in the database. Furthermore, within this paper, we also introduce the intended methodology towards incorporating the results of the qualitative and quantitative performance tests for the incremental improvement of drone operator training procedures.

The developed tools for video quality analysis and pilot performance modelling have important applications, as they can be further used for providing automated assistance to the drone pilot or for providing real-time qualitative feedback on flight performance. For this purpose, we present an AI drone copilot, building on the work in [44] that enables to detect 'bad' flight behaviour and warn the pilot in real-time. Furthermore, building on the algorithmic approach for video quality assessment [37], an automated tool is developed that enables to calculate optimal drone trajectories for target observation. This enables drone pilots to simply flip a lever on their remote control to obtain an optimal depiction of their target. Together, these automated assistance tools provide significant novel tools benefiting both novel and experienced pilots.

3. Virtual Environment for Quantitative Assessment

3.1. Software Framework

The proposed virtual environment for quantitative assessment is called the ALPHONSE simulator and consists of an interplay between multiple system components, as depicted on the system architecture diagram of Figure 3. The different system components are described in the following subsections. All the different system components communicate with one another over TCP and UDP connections. This implies that they can also be spread out over multiple computers, in order to spread the processing burden. In practice, we did not do this for the implementation presented in this paper,

which does imply that a reasonably powerful PC is required in order to run the simulator at maximum image resolution, while maintaining a low-latency control loop.

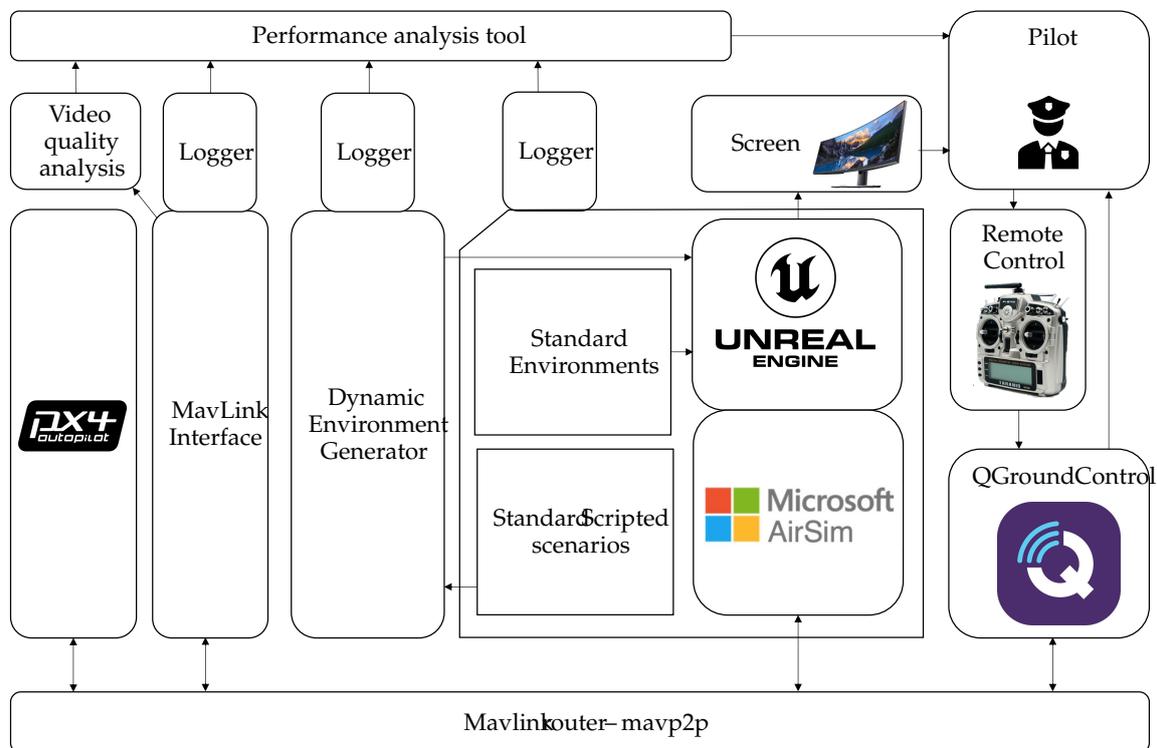


Figure 3. Software architecture of the proposed drone simulator system.

3.1.1. PX4

PX4 is an open-source flight control software for drones and other unmanned vehicles [71], offering a suite of tools for drone developers to share technologies and create solutions for various drone applications. PX4 establishes a standard for delivering drone hardware support and software stacks, facilitating a scalable ecosystem for building and maintaining both hardware and software.

Simulators enable PX4 flight code to control a computer-modelled vehicle in a simulated environment. Users can interact with this vehicle as they would with a real one, using QGroundControl as a Ground Control Station application or a radio controller. PX4 supports both Software In the Loop (SITL) simulation, where the flight stack runs on a computer, and Hardware In the Loop (HITL) simulation, which uses simulation firmware on a real flight controller board.

In our simulation framework, we utilize the version v1.10.1 PX4 SITL as the vehicle simulator. Given that PX4 is a widely adopted flight controller, this allows us to easily change the vehicle model if required. The SITL implementation of the PX4 autopilot provides a highly accurate simulation of the vehicle response, as it is a software version of a real autopilot hardware.

PX4 employs a simulation-specific module to listen on TCP or UDP ports, enabling simulators to connect to these ports and exchange information using the MavLink protocol [75]. MAVLink is a very lightweight messaging protocol for communicating with drones, used by a wide range of drones and drone users.

3.1.2. Mavlinkrouter

In order to pass the MavLink messages from one simulator component to another, we use the mavp2p MavLink router. This is a versatile and efficient Mavlink proxy, bridge, and router, implemented as a command-line utility [76]. It is primarily utilized to connect UAV flight controllers, interfaced via serial ports, with ground stations over a network. Additionally, mavp2p can facilitate

routing across various configurations involving serial, TCP, and UDP connections, thereby enabling communication across different physical and transport layers.

3.1.3. Mavlink Interface

This Python v3.8 program monitors messages transmitted by the PX4 SITL. It simultaneously stores these messages and provides real-time flight performance analysis. The MavLinkInterface program is implemented as a command-line tool and performs several critical functions. It connects to and arms the vehicle, then listens for various messages, including drone attitude, global position, local position, system status, vibrations, head-up display parameters, RC button status, RC raw channels, RC manual commands, and pilot home position. Additionally, it continuously calculates the distance and bearing between the pilot and the drone (in order to be able to simulate RF signal degradation over distance) and saves all the aforementioned data to a CSV file. Running the MavLinkInterface generates a CSV file of approximately 1 GB per minute, facilitating detailed post-flight analysis and data archiving.

3.1.4. Standard Scripted Scenarios

The Alphonse simulator is open and modular for including a wide range of potential scenarios. However, for this paper, we focus on one specific use case. This use case is defined by the Belgian Defence Special Forces Group and consists of an ISR (Intelligence, Surveillance, Reconnaissance) mission in a mountainous environment. The task of the drone pilot is dual. On the one hand, the pilot needs to detect an enemy encampment in a very large area, providing a high-quality video of this target and an accurate geo-location. On the other hand, the pilot also needs to detect and identify enemies present in this environment, while avoiding detection. When detecting any of these targets, the drone pilot is requested to flip a switch on the Remote Control to record the detection. This operation takes place in a complex environment, consisting of mountains, roads, lakes, and under varying weather conditions, in order to add complexity, such that pilot performance under stressful conditions can be assessed. The implemented scenario consists of enemy 1 camp which is statically allocated at run-time and 100 enemies that are dynamically allocated at run-time. Within the attributed time of 25 minutes that the pilots can operate in the simulator (before the battery of the drone runs out), it is quite impossible to detect all enemies in this environment, which creates extra pressure for the pilots.

3.1.5. Unreal Engine

The Unreal Engine, developed by Epic Games, is a state-of-the-art, real-time 3D creation tool widely recognized for its versatility and high fidelity in rendering complex, interactive environments [9]. Initially designed for video game development, the Unreal Engine has evolved into a powerful platform with applications extending far beyond gaming, including architectural visualization, film production, virtual reality, augmented reality, and scientific simulations. Its robust framework, supported by advanced features such as photorealistic rendering, dynamic lighting, and physics-based simulation, enables researchers and developers to create highly realistic and immersive virtual environments. The engine's extensive scripting capabilities, facilitated by its visual programming language, Blueprints, and support for C++, provide users with the flexibility to develop customized solutions for diverse scientific and industrial applications. The Unreal Engine's ability to handle large datasets and complex simulations makes it an invaluable tool for conducting high-fidelity virtual experiments and developing innovative visualization techniques in various scientific domains.

We used version 4.25 of the Unreal engine and implemented the following capabilities in order to fulfil the requirements of the simulator:

- Next to the standard First-Person-View camera, we introduce a ground-based observer viewpoint camera which renders the environment through the eyes of the pilot. At the start of the operation, the simulator starts in this ground-based observer viewpoint. As the environment is very large

and as most operations will take place Beyond Visual Line Of Sight (BVLOS), the operator will after a while usually switch to First-Person-View.

- Measure the number of collisions with the environment, as this is a parameter for the pilot performance assessment
- Measure the geo-location of any target (static camp or dynamic enemies)
- Calculate at any moment the minimum distance to any enemy
- Calculate the detectability of the drone, taking into consideration a detectability model in function of the distance of the drone to the enemy and the noise model of the drone type used [77,78].
- Sound an alarm when the drone is detected
- Include a battery depletion timer (set at 25 min). The drone crashes if it is not landed within the set time.

The Unreal engine renders photorealistic environments and can do that also directly in a Virtual Reality interface. However, while the simulation engine supports virtual reality and we have the equipment available, we have especially opted for using a curved monitor and *not* making use of a virtual reality interface for two reasons:

1. We want to avoid measuring the side-effects of virtual embodiment, where some pilots may be subject to.
2. Virtual reality would obstruct the use of exteroceptive sensing tools for measuring the physiological state of the pilot during the test.

3.1.6. Standard Test Environment

At the core of the standard test environment, we use the LandscapeMountains outdoor environment, which is available as a free download from the Unreal marketplace as a showcase of some advanced lighting, water reflection and weather effects. This environment features weather effects such as clouds and fog and also includes dynamic aerial obstacles (also called birds), thereby providing a rich training environment for operating drones. It is a very large (6GB in size) and excellently suited for testing outdoor surveillance scenarios in vast open space environments. Figure 4 gives the reader a general view of the scenery and the different environmental features of the environment.

Within this environment, we introduce two standardised targets, conforming to the ISR scenario:

- **A standardised visual acuity object.** It consists of a mannequin with a letter written on its chest plate, as shown on Figure 4e. The user will be requested to read out the letter (like for an ophthalmologist exam). The letters can be dynamically changed for every simulation and the mannequins can be spread randomly over the environment. In order to increase the level of difficulty, some of the mannequins are placed inside enclosures, forcing legibility only under certain viewing angles, thereby making the job even more difficult for the drone pilot.
- **An enemy camp.** It consists of a series of a series of tents, military installations and guarded watchtowers, as shown on Figure 4f.

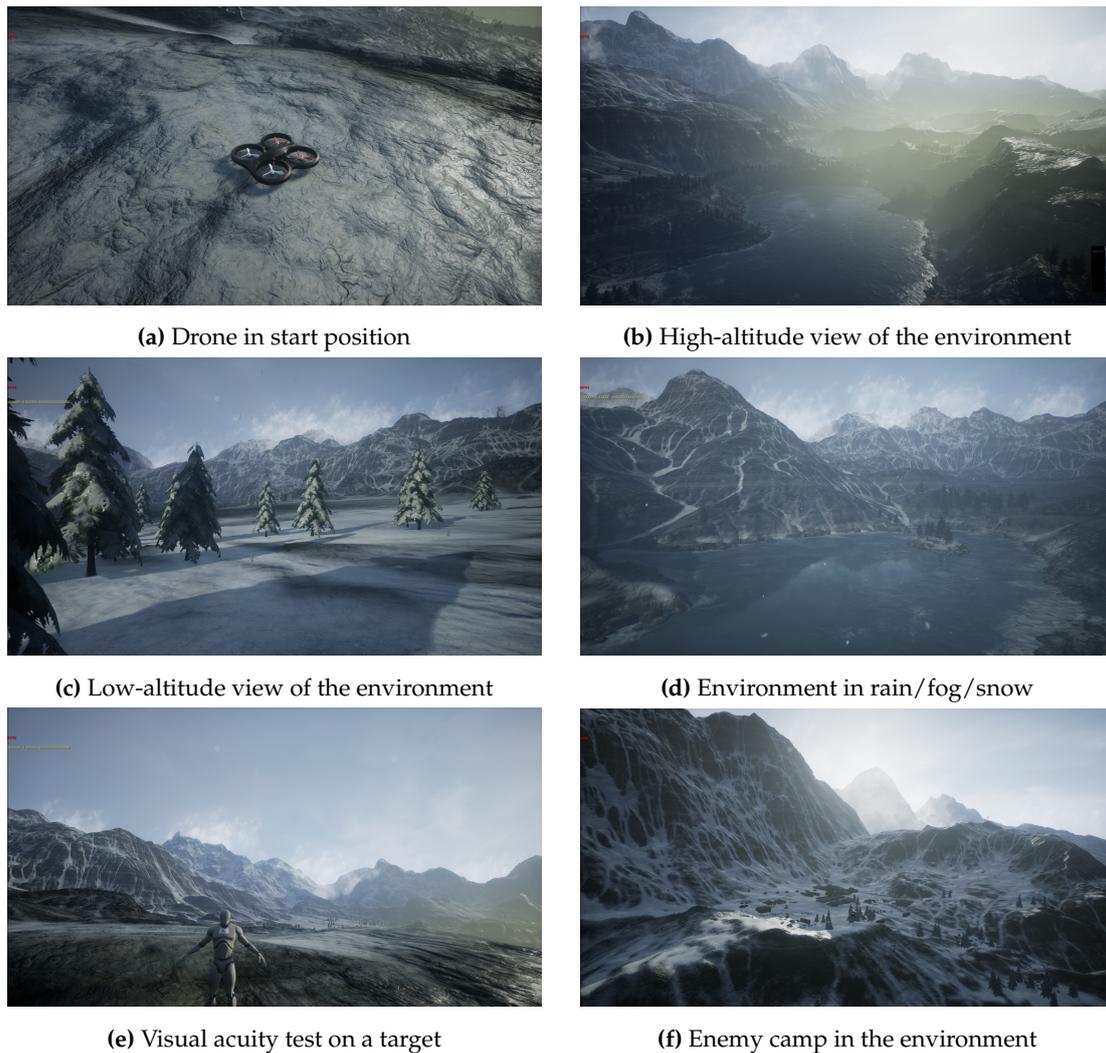


Figure 4. Snapshots of the simulator environment, showing the elements of the standardised test environment

3.1.7. AirSim

Microsoft AirSim is an advanced, open-source platform designed for the simulation of autonomous vehicles and drones [8]. Developed by Microsoft Research, AirSim provides a highly realistic, physics-based simulation environment that enables researchers and developers to train, test, and validate autonomous systems in a controlled, virtual setting before deploying them in the real world.

AirSim is open-source, cross platform, and supports hardware-in-loop with popular flight controllers such as PX4 for physically and visually realistic simulations. It is developed as an Unreal plugin that can be dropped into any Unreal environment. Conceptually, one could say that AirSim bridges the physical simulation of the environment (by the Unreal Engine) and the physical simulation of the drone itself (by the PX4 flight controller).

3.1.8. Dynamic Environment Generator

One of the primary objectives of the Alphonse flight simulator is to evaluate drone pilot performance under various external factors that may influence their flight capabilities. Two significant sources of disturbance are weather conditions and auditory disruptions. To simulate these disturbances, the Dynamic Environment Generator (DEG), a Python script, dynamically generates such environmental variations.

As outlined in Section 3.1.6, the standard test environment incorporates weather effects such as rain, snow, clouds, wind, and fog, as depicted in Figure 4d. The DEG introduces these environmental conditions at random intervals throughout the simulation runs to evaluate pilot performance under diverse weather scenarios.

Additionally, to assess pilots' resilience to auditory disturbances, the DEG cycles through various auditory tracks during the simulation runs. This approach allows for a comprehensive assessment of pilot performance under different types of auditory disruptions.

3.1.9. QGroundControl

QGroundControl is an advanced ground control station for UAVs [80], offering comprehensive flight control and mission planning capabilities for any MAVLink-enabled drone, as well as vehicle setup for PX4 autopilot-powered UAVs. The software is entirely open-source and is used by a wide range of drone users, which makes it the logical choice as a software platform for planning and controlling the flight operations, as many of our end-users will already be used to this tool.

In our simulation framework, we use QGroundControl for planning missions (e.g. setting geofences, defining mapping tasks, defining waypoints or regions of interest, etc.) and for arming and disarming the drone. All this is particularly useful, as many of our operators have indicated to work in Beyond Visual Line of Sight (BVLOS) conditions, where the use of tools like QGroundControl becomes essential.

In our application, v4.1.1 of QGroundControl is utilized to provide a graphical user interface (GUI) featuring an earth map, which aids in drone localization, as shown on Figure 5. Although these earth terrain maps do not automatically correspond to the simulated environments, integrating a reasonably accurate environmental model of a real-world setting allows the simulation and GUI to function in complete cohesion. This configuration provides an optimal training environment for Beyond Visual Line of Sight (BVLOS) flights.

Furthermore, QGroundControl supports input from a Remote Control connected via USB to the simulator PC. The Remote Control can be calibrated and configured through QGroundControl to operate in various modes, tailored to the user's preferences. Next to the normal mapping of flight controls, we also programmed specific levers on the Remote Control, such that when they are activated, all relevant flight data (position etc.) is logged in a separate CSV file. This enables the pilots to fully focus on flying during the simulation, as they can keep their both hands on the Remote Control at all times.

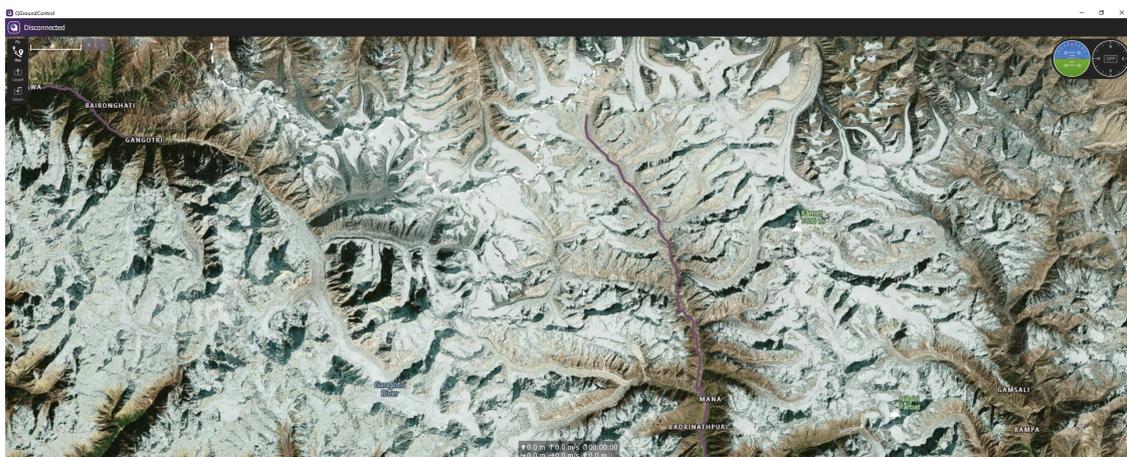


Figure 5. Interface of the QGroundControl application.

3.1.10. Logging Systems

Attached to the various components of the drone simulator software are multiple logging systems that enable off-line processing by the performance analysis tool. In general, three types of logs are created:

1. The MavLink interface records a series of interesting parameters related to the drone itself (its position, velocity and acceleration, control parameters, vibrations, ...) on the MavLink datastream.
2. The Dynamic Environment Generator records the environmental conditions (wind direction and speed, density of rain, snow and fog, ...) and the presence of auditory disruptions (type of audio track, sound intensity, ...).
3. The Unreal-based simulation records the number of collisions with the environment, the geo-location of all targets, the minimum distance to any enemy and the detectability of the drone.

Next to these three inputs, the video quality analysis tool, described in section 3.2 also provides an important input to the performance analysis tool.

3.1.11. Video Quality Analysis Tool - See Section 3.2

3.1.12. A Mission Performance Analysis Tool - See Section 4.1

3.2. Quantitative Evaluation of Drone-Based Video Quality

3.2.1. Concept

As established in Section 1, there is currently no tool available to determine whether a video produced by a specific drone operator contains sufficient information about a particular target. This paper proposes a methodology to quantitatively assess the content of drone-based video data.

It is important to note that this methodology does not rely on video signal analysis, as such an approach would be challenging to adapt across various applications or mission scenarios. Instead, our methodology is based on the analysis of positional data, which drones typically acquire via their positioning sensors. This approach is inherently task-agnostic and can be applied to a wide range of applications.

For this paper, we primarily focus on military operations, where the objective is to gather the maximum amount of data about a target in the minimum amount of time. A limitation of this application choice is that the proposed methodology does not account for cinematographic constraints (e.g., the rule of thirds) commonly used in professional video photography, thus limiting its applicability to such contexts.

3.2.2. Methodology

The video quality analysis methodology presented here is designed to be as task-agnostic as possible. However, certain key basic assumptions must be defined for the algorithm:

- We assume that the drone camera is always directed at the target. This assumption simplifies the algorithm by avoiding (dynamic) viewpoint adjustments based on drone movement. This is a realistic scenario, as in actual operations, a separate camera gimbal operator typically ensures the camera remains focused on the target. This task can also be automated using visual servoing methodologies [67], which we assume to be implemented in this paper.
- To ensure uniform perception of the target from various viewing angles, we assume the target has a perfect spherical shape. While this is an approximation and may differ for targets with non-spherical shapes, it is the most generic assumption and can be refined if specific target shapes are more applicable to particular uses.
- Since the zoom factor is not dynamically available to the algorithm, we assume a static zoom factor.

- The input parameters for the video quality assessment algorithm are the drone's position at a given time instance $\mathbf{x}_i = (x_i, y_i, z_i)$ and the target's position $\mathbf{x}_t = (x_t, y_t, z_t)$, which is assumed to remain static throughout the video sequence.

The proposed methodology for quantitative video quality analysis considers three sub-criteria that together determine the overall measure of video quality. These metrics are:

1. *The number of pixels on target ϕ_p .* It is well-known that for machine vision image interpretation algorithms (e.g., human detection [68], vessel detection [69]), the number of pixels on target is crucial for predicting the success of the image interpretation algorithm. Similarly, for human image interpretation, Johnson's criteria [70] indicate that the ability of human observers to perform visual tasks (detection, recognition, identification) depends on the image resolution on the target. Given a constant zoom factor, the number of pixels on target is inversely proportional to the distance between the drone and the target, such that:

$$\phi_p = \frac{\lambda}{|\mathbf{x}_i \mathbf{x}_t|}, \quad (1)$$

where λ is a constant parameter ensuring that $0 \leq \phi_p \leq 1$, dependent on the minimum distance between the drone and the target, the camera resolution, and the focal length.

2. *The data innovation ϕ_d .* As discussed in the introduction, assessing the capability of drone operators to obtain maximum information about a target in minimal time is crucial. The data innovation metric evaluates the quality of new video data. This is achieved by maintaining a viewpoint history memory θ_j with $j = 1 \dots i - 1$, which stores all normalized incident angles of previous viewpoints. The current incident angle θ_i is compared to this memory by calculating the norm of the difference between the current and previous incident angles. The data innovation is the smallest of these norms, representing the distance to the closest viewpoint on a unit sphere:

$$\phi_d = \min_{j=1}^{i-1} (|\theta_i - \theta_j|) \quad (2)$$

New viewpoints should be as distinct as possible from existing ones, as expressed by (2).

3. *The trajectory smoothness ϕ_t .* High-quality video requires a smooth drone trajectory over time. Irregular motion patterns make the video signal difficult to interpret by human operators or machine vision algorithms. The metric ϕ_t evaluates trajectory smoothness by maintaining a velocity profile $\dot{\mathbf{x}}_j$ with $j = 1 \dots i - 1$, which stores all previous velocities. The current velocity $\dot{\mathbf{x}}_i$ is compared to the n most recent velocities. The norm of the difference between the current and previous velocities is weighted by recency. The weighted sum of the n most recent velocity differences measures changes in the motion profile and is inversely proportional to trajectory smoothness:

$$\phi_t = \frac{1}{\sum_{j=i-n}^{i-1} \frac{1}{i-n} |\dot{\mathbf{x}}_j - \dot{\mathbf{x}}_i|} \quad (3)$$

All three video quality sub-criteria produce values between 0 and 1. Assigning equal importance to each sub-criterion, the overall measure for drone-based video quality is:

$$\phi = |(\phi_p, \phi_d, \phi_t)| \quad (4)$$

Weights can be applied to this global metric to prioritize specific sub-criteria based on application requirements. However, this paper examines the generic case without applying any such weights.

3.2.3. Validation

For the validation of the proposed methodologies, we assessed the performance of drone operators in the simulation environment introduced earlier. Multiple operators with known proficiency profiles were tasked with producing a high-quality video of a target within the simulation environment, and the resulting total ϕ scores they obtained were recorded, as shown in Figure 6. The figure demonstrates that the algorithm effectively distinguishes between proficient users (e.g., operator 3) and less proficient users. However, further research is necessary to validate the relationship between subjective quality assessments and this objective metric.

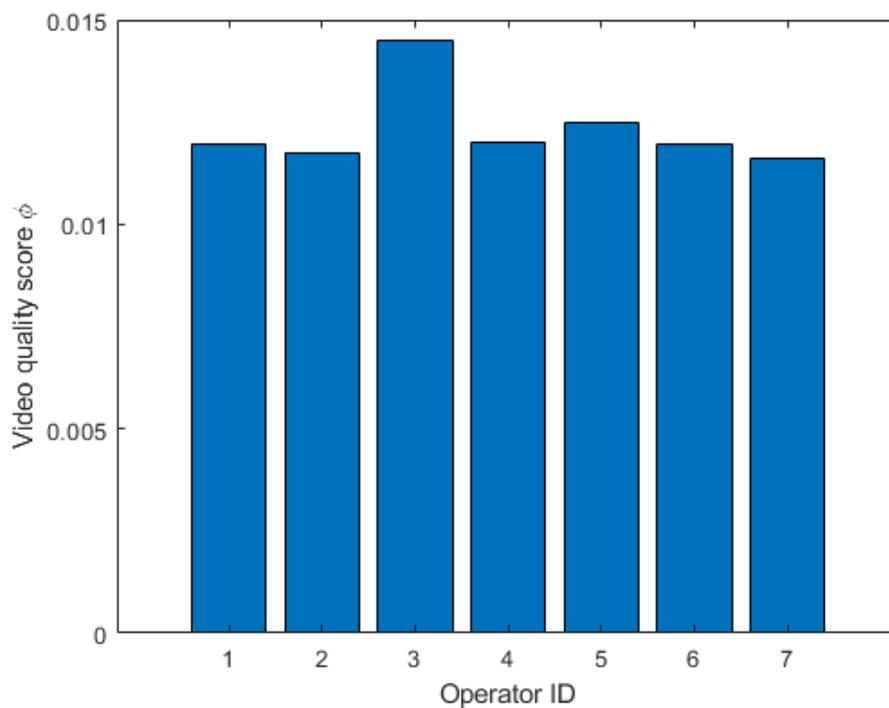


Figure 6. Video quality scores obtained by seven operators.

4. Drone Operator Performance modelling

4.1. Performance Analysis Tool

4.1.1. Metrics Definition

The role of the mission performance analysis component is to read in all the log files that are created after a simulator run by a pilot (which easily consists of a few 100GB of data) and present this data in a user-friendly way. For that reason, it extracts a number of performance metrics and visualizes those on a Graphical User Interface.

The most important performance metric is the performance score ψ that is defined as an average of multiple partial performance components:

$$\psi = \sqrt{\sum_{j=1}^6 \psi_j} \quad (5)$$

The first performance component analyses the control commands and assesses the smoothness of these pilot controls:

$$\psi_1 = \frac{\omega_1}{\prod_{i=1}^n \nabla v_i}, \quad (6)$$

where ω_1 is a weight factor and v_i consist of a number of ($n = 20$) control signals that can be read from the MavLink protocol. After a number of experiments, we concluded that the following combination of MavLink control signals offer the most relevant measure to assess the performance of the pilot:

- GPS latitude, altitude and heading
- Velocity in the X, Y, Z direction
- Roll, pitch and yaw angles
- Velocity in roll, pitch and yaw
- Throttle level
- Climb rate
- Vibrations in the X, Y, Z direction
- Control stick position in the X, Y, Z and roll direction

Note that equation 6 considers the gradient of the v_i signals, as it is not really their magnitude which is of importance, but more their smoothness.

The ψ_2 component assesses the number of collisions made by the pilot, as it is inversely proportional to the number of collisions with the environment κ as measured during the UnReal simulator during the execution of the flight:

$$\psi_2 = \frac{\omega_2}{1 + \kappa}, \quad (7)$$

The third component ψ_3 is quite specific for this use case scenario, as it assesses whether the pilot has succeeded in locating the enemy base. When the pilot triggers the detection reporting button on the remote control within a predefined distance that is close enough to the centre of the enemy camp, the boolean variable γ is set to one (otherwise it is zero). At the same time, the distance to the camp δ_C is measured, which serves as an error measure for the pilot reporting. In case that the camp was not found, this distance is initialised to a very high value. This enables to define the camp detection component as:

$$\psi_3 = \omega_3 \frac{\gamma}{\delta_C}, \quad (8)$$

The fourth component measures quality of the video recorded by the pilot, following the methodology developed in section 3.2 and, with ϕ as defined by equation 4.

$$\psi_4 = \omega_4 \phi, \quad (9)$$

The fifth component ψ_5 is a metric to assess target identification accuracy, as it compares the number targets that are identified τ_I with the total number of targets τ_N . The metric also takes into consideration the average distance to the detected targets τ_E and the total flight time T .

$$\psi_5 = \omega_5 \frac{\tau_I}{\tau_N \tau_E T}, \quad (10)$$

The final component ψ_6 further assesses the pilot performance in avoiding detection by enemies, by comparing the mean distance to enemies δ_E and the minimum distance to enemies to the total number of times the drone has been spotted by the enemies ϵ , as it was flying too close.

$$\psi_6 = \omega_6 \frac{\frac{1}{n} \sum_{m=1}^n \delta_E \min \delta_E}{1 + \epsilon}, \quad (11)$$

4.1.2. Interface Design

The mission performance analysis component was developed as a Matlab application, in order to make it user friendly for the analyst. The interface, as shown on Figure 7, features a large 2D top-down view of the standardised test environment, showing the locations of the different enemies (blue dots), the enemy camp (red square), and the drone trajectory (blue line). This allows an analyst to discuss with the pilot the performance in terms of trajectory smoothness and the adherence to certain predefined search patterns.

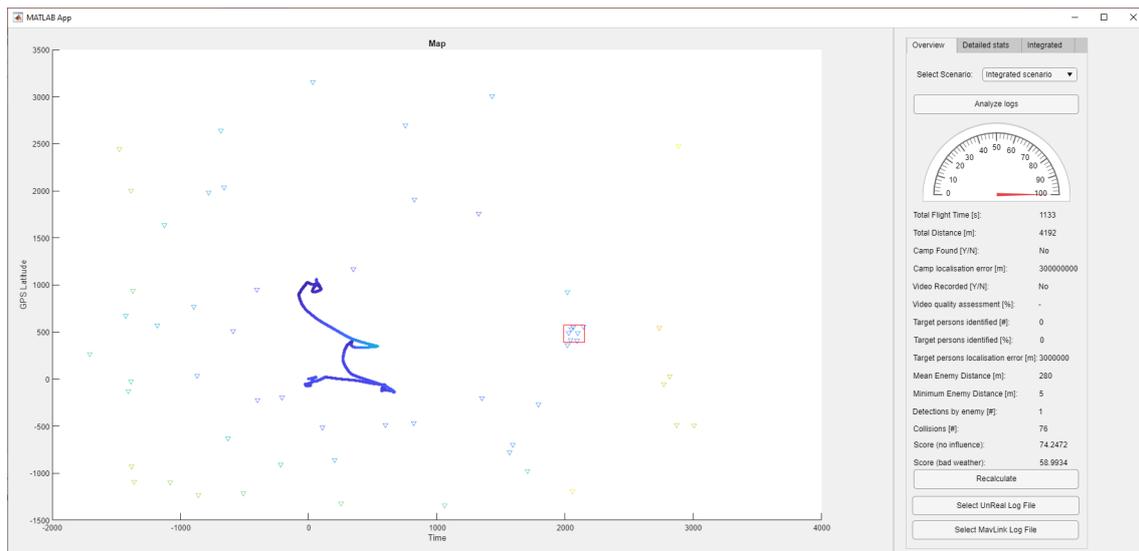


Figure 7. Main interface of the performance analysis tool, featuring on the right side a number of performance metrics and on the left side a top-down view of the standardised simulation environment, showing the locations of the different enemies (blue dots), the enemy camp (red square), and the drone trajectory (blue line).

The right side of the interface shows a number of metrics that can help for the analysis of the flight performance:

- The total flight time in seconds;
- The total distance flown in meters;
- Whether the camp has been found, and - if yes- the error on the distance measurement;
- Whether a video was recorded, and - if yes- the video quality score according to equation 4;
- The number and percentage of enemies identified and their localisation error;
- The mean and minimum enemy distance in meters;
- Whether the drone has been detected by enemies;
- The number of collisions;
- The performance score under multiple environmental; conditions or under multiple human factors. Note that - as can be expected - the performance score in normal weather is better than the performance score in bad weather.

The Interface also features a second screen that gives access to much more detailed performance data of 60 time signals over the whole flight duration, as shown in Figures 8 and 9. The reason for including this interface is that we wanted to provide a full explainability towards the singular performance score metric, thereby increasing the trust of the users in the system [72]. Indeed, by analysing the different time signals recorded during the flight, an expert analyst can deduce e.g. what went wrong during a certain flight and why the corresponding performance score is low.

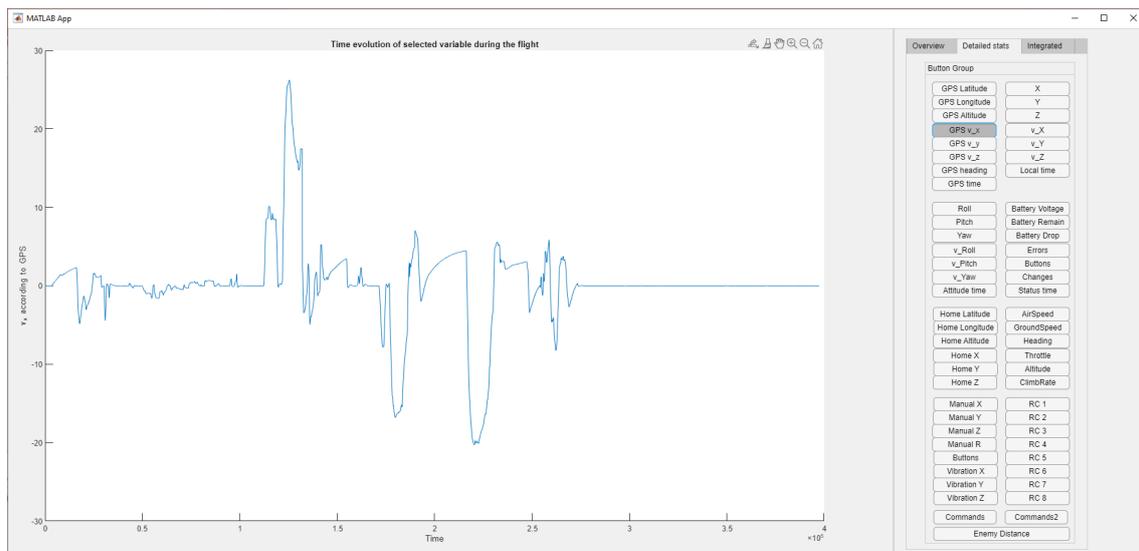


Figure 8. Expert analysis panel of the performance analysis tool. The user can select any of the 60 recorded time-signals on the right to better understand the composition of the performance score. Here, the graph shows the X-axis component of the drone velocity, as measured by its on-board GPS system.

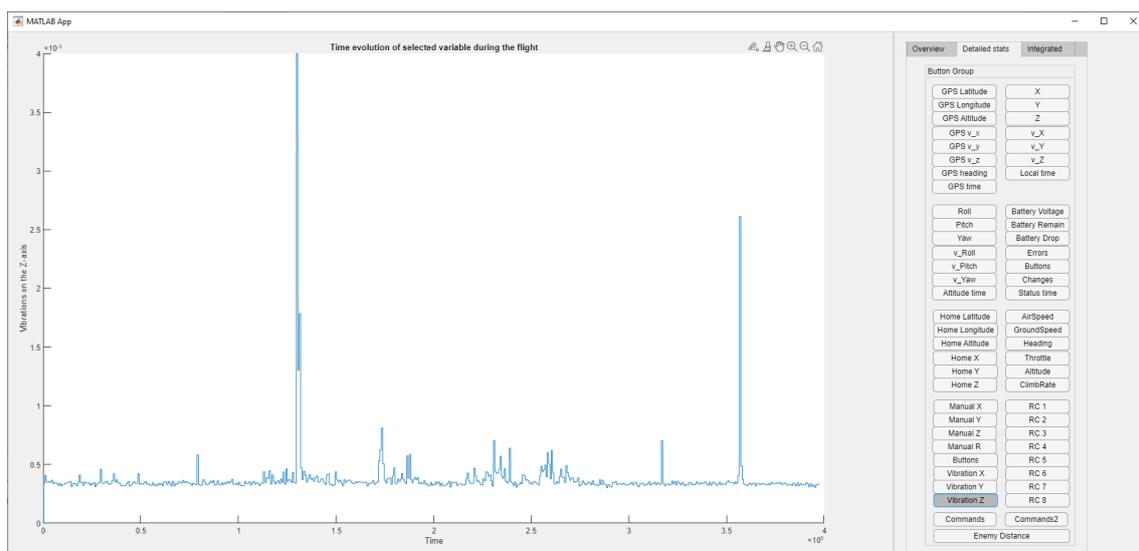


Figure 9. Expert analysis panel of the performance analysis tool. The user can select any of the 60 recorded time-signals on the right to better understand the composition of the performance score. Here, the graph shows the Z-axis component of the drone vibrations.

As an example, Figure 8 shows the X-axis component of the drone velocity, as measured by its on-board GPS system. The steep inclines in this graph do indicate erratic flight behaviour that will be penalised in the performance score.

Similarly, Figure 9 shows the Z-axis component of the drone vibrations. While the signal is mostly flat, some important spikes can be noted that likely indicate collisions with the environment, leading to a lower performance score.

4.2. Human Performance Model

4.2.1. Methodology

Human performance modelling involves creating a mathematical representation of human perception-reaction behaviours and cognitive reasoning processes. The objective of this modelling is to

enhance the safety, efficiency, and performance of human-machine systems, such as drone pilot-drone interactions.

We have developed an innovative drone operator performance model, particularly focusing on military operations. To collect comprehensive user inputs, standardized questionnaires were designed to encapsulate pilot performance within a human performance model that considers a broad set of parameters. This model assesses the impact of varying operator training and skill levels, as well as the diverse capabilities and reliability of different platforms.

The proposed human performance model enables the examination of multiple system prototypes and theories related to human performance, which can be further validated through human-in-the-loop experiments. This adaptability is crucial for keeping pace with rapidly evolving drone technologies and for accurately representing human behaviour under extreme conditions. Objective functions are constructed to optimize the model's output—such as error rate, task completion time, and workload—by treating model inputs as decision variables.

To evaluate the relationship between human factors and operator performance, we employed a user-centred design approach [73]. Following this methodology, we identified human factors that potentially impact drone pilot performance through interviews with experienced drone operators. The key factors, determined by expert operatives, are listed in Table 1. Additional factors, such as distraction, were also considered but scored low and were excluded from the primary list. These identified parameters are re-evaluated with test subjects during an intake questionnaire to assess their state before the simulation exercise.

Table 1. Most important human factors impacting drone operator performance.

Human Factor	Importance level (0 – 100%)
Task Difficulty	89%
Pilot Position	83%
Pilot Stress	83%
Pilot Fatigue	83%
Pressure	83%
Pilot subjected to water or humidity	83%
Pilot subjected to temperature changes	78%
Information location & organization & formatting of the controller display	78%
Task Complexity	78%
Task Duration	78%
Pilot subjected to low quality breathing air	72%
Pilot subjected to small body clearance	72%
Ease-of-use of the controller	72%
Pilot subjected to noise / dust / vibrations	67%
Task Type	67%

Secondly, we identified operational scenarios and environmental conditions affecting drone pilot performance through further interviews with security sector operators. This led to the compilation of standard operational scenarios that cater to diverse end-user needs, including complex target observation and identification missions in both urban and rural environments.

Thirdly, we developed a simulation environment for complex drone operations, as presented in section 3. Within this simulation environment, pilots face dynamic environments, changing weather conditions, and time pressure, all of which can induce errors impacting performance. This simulation environment is completely open and customizable, which enables us to incorporate the standard test scenarios, multiple customizable drones and to quantitatively measure the performance of the pilots on-line while executing the mission. Next to this interoceptive sensing of the human physiological state using the metrics presented in section 4.1.2, we also plan to use (in a later stage) exteroceptive sensing of the human physiological state by a camera system targeted at the pilot, estimating fatigue etc.

Following mission completion, test subjects complete an additional questionnaire to assess their physiological state and any changes since the intake survey. The collected data comprises human factors and physiological states before, during, and after the mission, alongside quantitative performance data from the simulation engine. This comprehensive dataset enables the development of a mathematical model correlating human factors and physiological states with performance. Such a model can predict performance based on given input states and has potential applications in drone pilot accreditation and certification processes.

4.2.2. Results & Discussion

To evaluate the performance model, we recruited a diverse group of pilots to undertake several missions using the proposed simulator system. This group comprised individuals with varying degrees of experience in drone operation. The test group includes both military and civilian personnel from Belgian Defence, with the military staff consisting mostly exclusively of drone pilots from the Special Forces Group. While these are highly trained individuals with experience in operating military drones, they do not necessarily all have specific type training on quadrotor drones. The civilians included in the test panel included researchers working at the Belgian Royal Military Academy with highly varying skill levels related to drone piloting.

The initial goal was to subject the test pilots to a whole range of human factors, as given by Table 1. However, this research study was not performed under the umbrella of an ethical committee, so it was not possible to perform research on humans and to subject the human test subjects to a series of human factors like sleep deprivation, water, humidity & temperature changes, etc. Therefore, we concentrated on a number of factors that we could readily measure with the simulation system.

The initial aspect we examined was the decline in pilot flight performance under deteriorating weather conditions. As previously mentioned, the simulator subjects pilots to a range of weather scenarios (wind, snow, rain, fog, etc.). Among these variables, wind is the most straightforward to quantify, as illustrated in Figure 10.

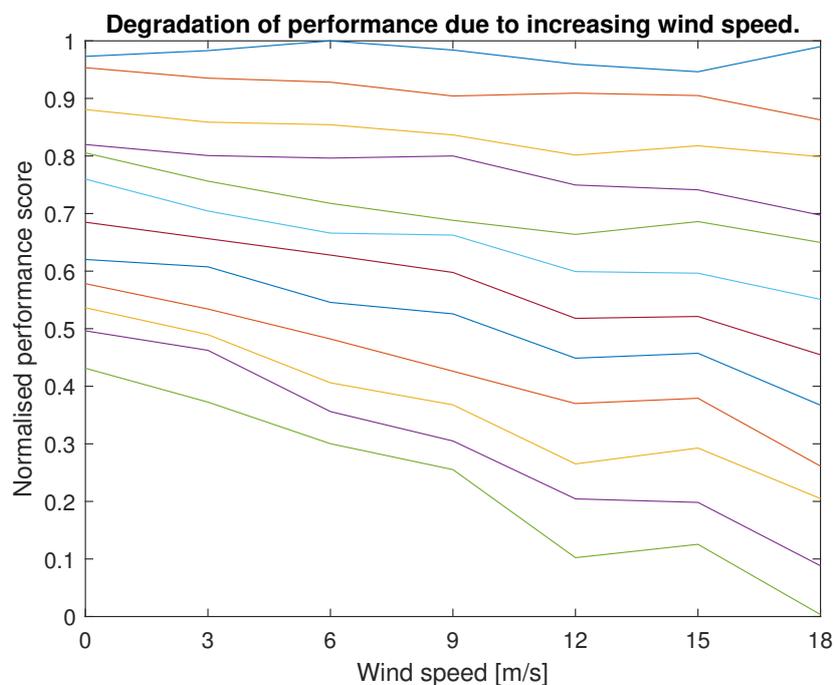


Figure 10. Degradation of performance due to increasing wind speed. Each line represents the performance score of an individual pilot across various wind speeds. The figure shows that good pilots manage to keep a consistent performance level, whereas bad pilots see their performance reduced.

In Figure 10, each line represents the performance score of an individual pilot across various wind speeds. As anticipated, the performance scores generally exhibit a downward trend, indicating that performance declines as wind speed increases. However, this trend should not be generalised, as Figure 10 also reveals that highly skilled pilots show only minimal performance degradation under increased wind speeds. This suggests that proficient pilots can effectively manage this external factor, whereas less skilled pilots experience significant performance decreases in high wind conditions.

Another aspect we modelled was the degradation of pilot flight performance due to auditory disturbances. The simulator facilitates this by introducing distracting noises during flight execution. The analysis of performance scores under auditory disturbances initially puzzled us, as the results were highly individualistic: some pilots managed various types of auditory disturbances exceptionally well, while others experienced decreased flight performance. However, when we cross-referenced these observations with the profiles of the test subjects, we discovered that military drone pilots were generally resistant to auditory disturbances, whereas civilian pilots exhibited reduced performance under such conditions, as illustrated in Figure 11.

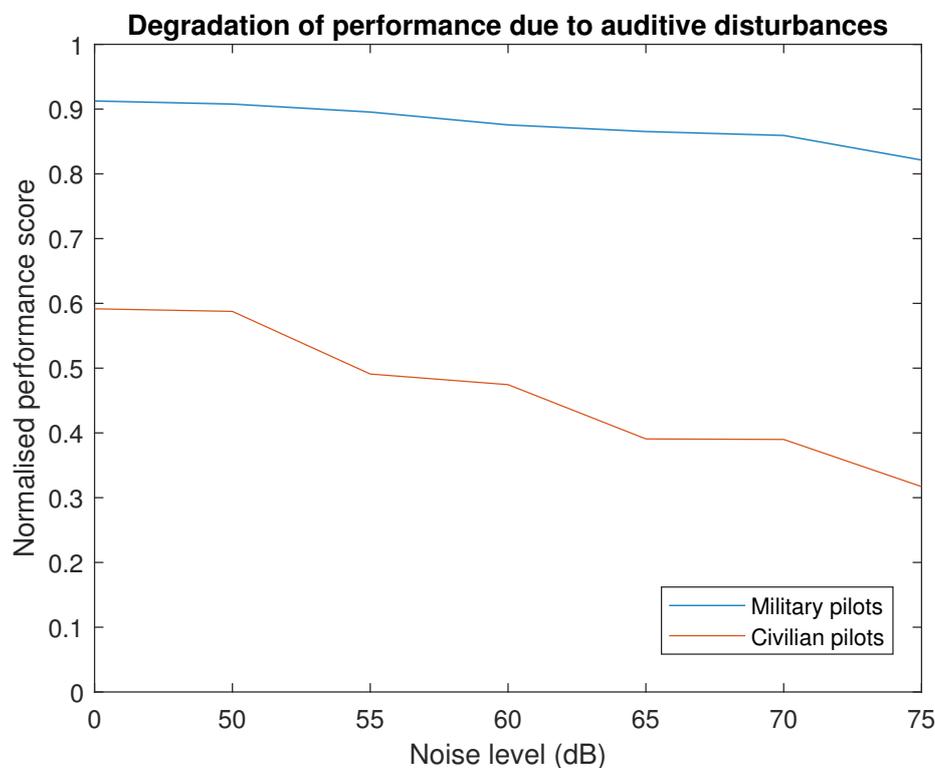


Figure 11. Degradation of performance due to auditive disturbances level. The figure shows that military pilots (blue line) handle auditive disturbances very well, whereas civilian pilots tend to have their performance reduced when subjected to auditive disturbances.

This observation is likely linked to the rigorous military training that prepares personnel to handle auditory disturbances. Notably, the group of military pilots comprised Special Forces personnel with operational experience, equipping them with the skills to overcome such disturbances. This background also explains their higher average performance level compared to civilian pilots.

Stress is a frequently reported factor influencing pilot flight performance. To quantitatively measure pilot stress levels, we included the Holmes-Rahe Life Stress Inventory [74] in the intake questionnaire, which provides a quantitative assessment of stress. The variation in performance scores relative to different amounts of stress levels is depicted in Figure 12.

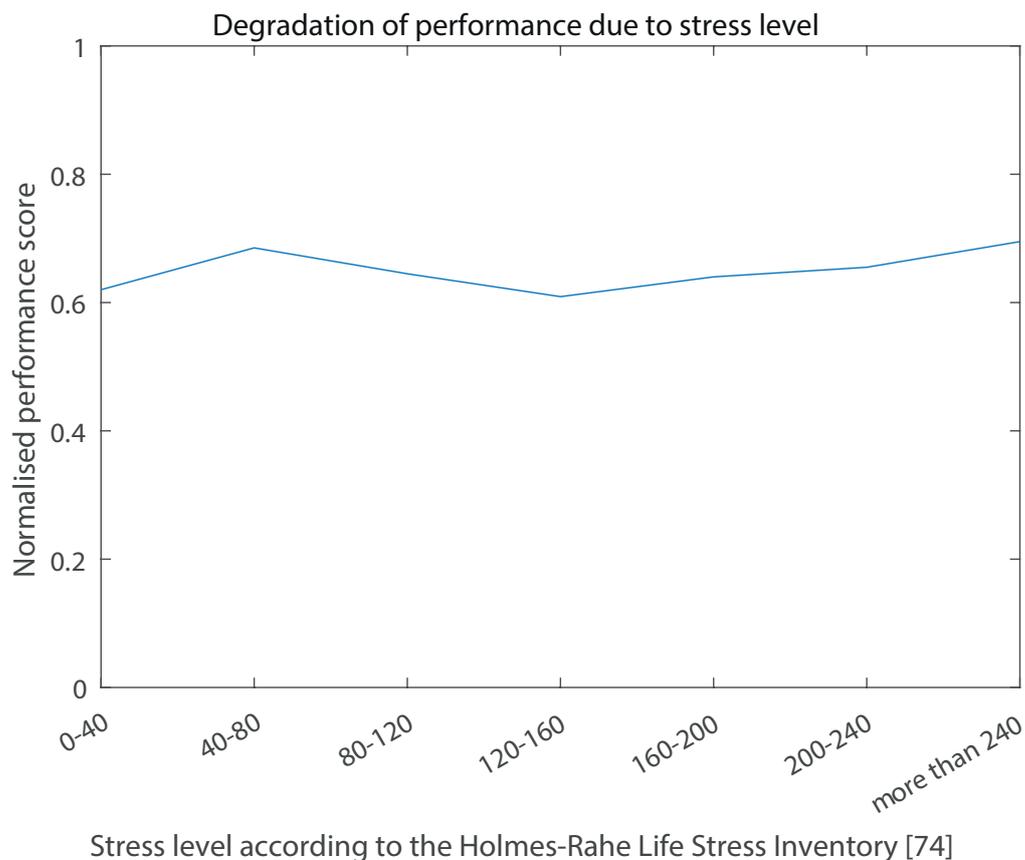


Figure 12. Degradation of performance due to a-priori stress level. No clear correlation between the a priori stress level and the pilot flight performance could be established.

To be completely honest, we were initially disappointed with the results shown in Figure 12, as it reveals no clear correlation between stress levels and flight performance. Based on existing literature, we anticipated a decreasing trend; however, our experiments did not confirm this. This discrepancy may stem from three cumulative factors

- Measurement of stress level only before the test and not during the test may not be enough to yield appropriate measurements of stress levels. As discussed before, we intend to extend the simulator system with exteroceptive sensing systems that would enable the measurement of stress levels during flight, but this is not something that is ready yet.
- The Holmes-Rahe Life Stress Inventory is likely not be the optimal tool for quantifying stress levels in this context, as it emphasizes long-term life events rather than short-term stressors.
- Limited sample size of our pilot population, which may not be sufficient to yield statistically significant results.

Another possible reason is that

A final correlation we aimed to quantify is the relationship between training and performance, as illustrated in Figure 13. This figure clearly demonstrates that to enhance performance, training must be specific to the type of drone being operated. In other words, cross-type training does not significantly improve performance, as evidenced by the blue line in Figure 13. This line predominantly represents performance data from pilots with substantial experience flying fixed-wing drone systems. However, their fixed-wing experience did not translate into high performance scores in the simulator. It is important to note that the Alphonse simulator and its performance metrics focus on assessing the ability to control the drone, rather than other critical aspects of drone operation such as ATM and

UTM interactions, RF communication, and weather assessment, areas where fixed-wing pilots would likely excel.

The correlation between flight training on quadrotors and performance scores is immediately evident, as shown by the red line in Figure 13. This curve displays a steady and monotonous increase, suggesting that more training consistently leads to better performance. The highest results in our simulator were indeed achieved by a Special Forces operative who flies quadrotors on a daily basis, thereby accumulating a substantial number of flight hours.

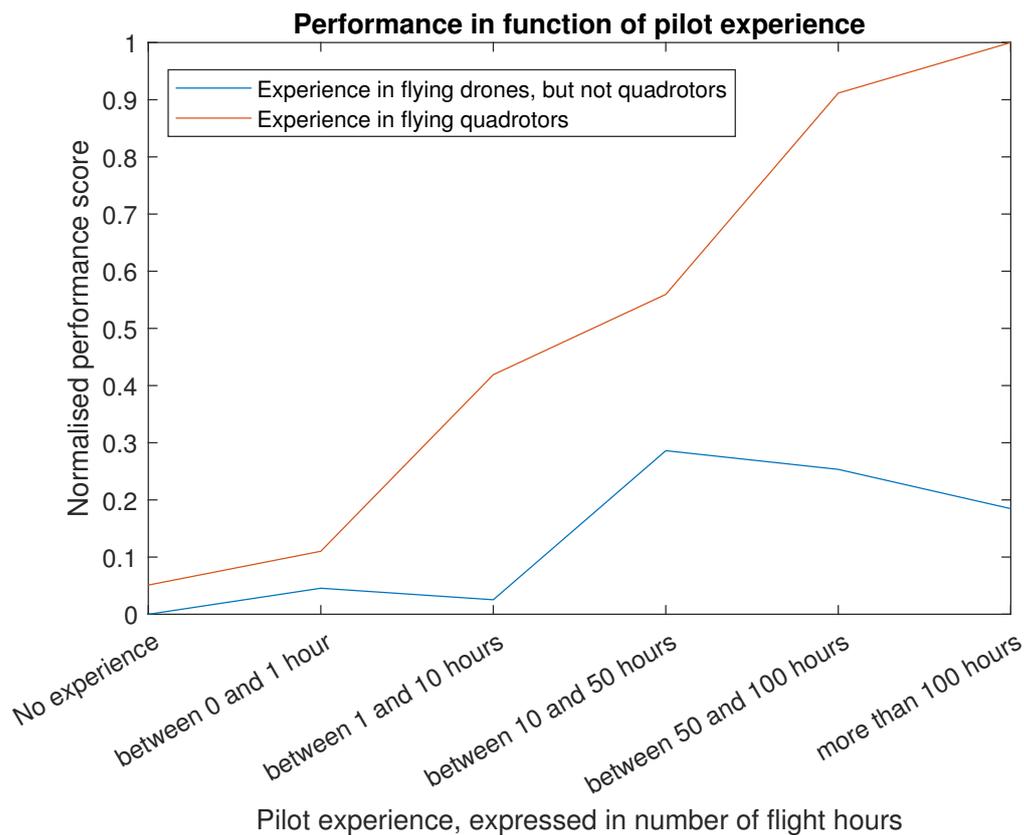


Figure 13. Degradation of performance due to pilot experience. The figure shows that training drastically improves performance (red line), but the training should be specific to the type of drone, as training on fixed wing drones (blue line) will not help much for increasing the performance on quadrotors.

5. Application Use Cases

5.1. AI Copilot for Drone Operator Assistance

5.1.1. Motivation and Concept

Given the susceptibility of drone operations to incidents arising from human error, it is imperative to provide robust support to human pilots. The automotive industry has already addressed similar issues through advanced driver assistance systems, which alert drivers when they are inattentive or when they deviate from their lanes without signalling. Drawing on the data accumulated from the pilot performance assessment system presented in this study, a comparable system for drone operations can be developed.

This system utilizes an artificial intelligence (AI) framework that analyses flight patterns across a spectrum of pilots, ranging from novices to highly trained experts. By examining these parameters, the AI system can identify behaviours characteristic of proficient pilots and those indicative of sub-

optimal performance. This paper details the development of an AI-based expert system designed to perform such assessments in real-time. The expert system functions as a virtual co-pilot, continuously monitoring flight performance and issuing warnings when detrimental flight behaviours are detected.

5.1.2. Methodology

In the ALPHONSE simulator, multiple drone pilots of varying skill levels performed flight trials under controlled conditions. Each pilot executed a multi-target detection mission while being subjected to auditory distractions, changing weather conditions, and other variables. Concurrently, 66 flight parameters available in the MavLink drone messaging protocol [75] were tracked at a rate of 300 samples per second. The primary challenge in designing the virtual co-pilot system is determining the relationship between these flight parameters and the pilot's skill level. To address this, a two-step approach was employed.

In the first step, termed the 'human intelligence' step, the problem space was narrowed using insights from the drone pilots, gathered through intake and out-take questionnaires. This process reduced the 66 flight parameters to a list of 26 relevant parameters for further analysis. These parameters include speed and acceleration, gradients for roll, pitch, yaw, and climb rate, vibrations, control stick inputs, among others.

In the second step, the 'artificial intelligence' step, a neural network was trained to model the relationship between the 26 flight parameters and pilot skill levels. Initially, data pre-processing was necessary due to the high sampling rate of the flight simulator logs. To manage the data flow, an averaging process over 100 samples was applied, equating to approximately 0.3 seconds per sample. The sequence length had to balance two competing requirements: maximizing data inclusion in the decision process while minimizing delay in skill recognition. A sequence length of 50 was selected, corresponding to around 15 seconds.

The neural network classifier comprises 100 hidden layers and five base layers:

1. A SequenceInputLayer that handles sequences of the 26 flight parameters.
2. A bidirectional long short-term memory layer that learns bidirectional long-term dependencies between time steps of time series.
3. A fully connected layer that multiplies the input by a weight matrix and adds a bias vector.
4. A softmax layer that applies a softmax function [81] to the input.
5. A classification layer that computes the cross-entropy loss for classification tasks with mutually exclusive classes.

The outputs were categorized into three classes:

- **Category 1:** Novice pilots, which also includes pilots experienced with fixed-wing drones but not rotary-wing drones, subject to poor skill transfer across drone types, as observed and discussed in section 4.2.2.
- **Category 2:** Competent pilots with experience flying rotary-wing drones.
- **Category 3:** Expert pilots, consisting of highly skilled pilots who regularly practice complex flight operations.

The Adaptive Moment Estimation (ADAM) method for stochastic optimization [82] was implemented. ADAM combines adaptive gradients and root mean square propagation, allowing the algorithm to converge to a stable solution for the classifier, as illustrated in Figure 14.

This approach demonstrates the efficacy of integrating human intelligence and artificial intelligence to enhance drone pilot training and performance assessment, ultimately contributing to the development of advanced support systems for unmanned aerial operations.

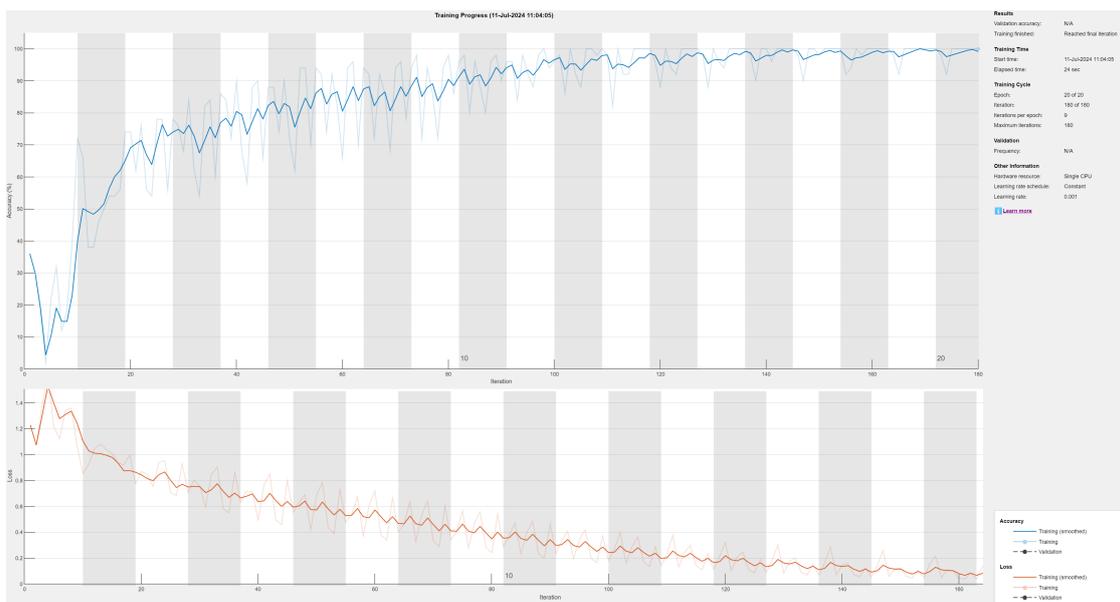


Figure 14. Convergence of the accuracy and loss of the ADAM pilot skill classifier.

5.1.3. Validation and Discussion

In this section, we present the results of tests conducted with drone pilots from Belgian Defence and civilian researchers, who have flown within the ALPHONSE simulator. These pilots initially provided flight data to train the model and subsequently participated in the model’s validation. For validation purposes, the available flight data from all pilots was divided, with 80% used for training and 20% reserved for validation.

To evaluate the performance of the pilot skill level classifier, the Receiver Operating Characteristic (ROC) curve for the classifier is shown on the left side of Figure 15. This ROC curve is zoomed in towards the upper left corner, highlighting the low ratio of false positives to true positives across all classes. The confusion matrix, displayed on the right of Figure 15, further indicates the classifier’s overall accuracy at 88.5%.

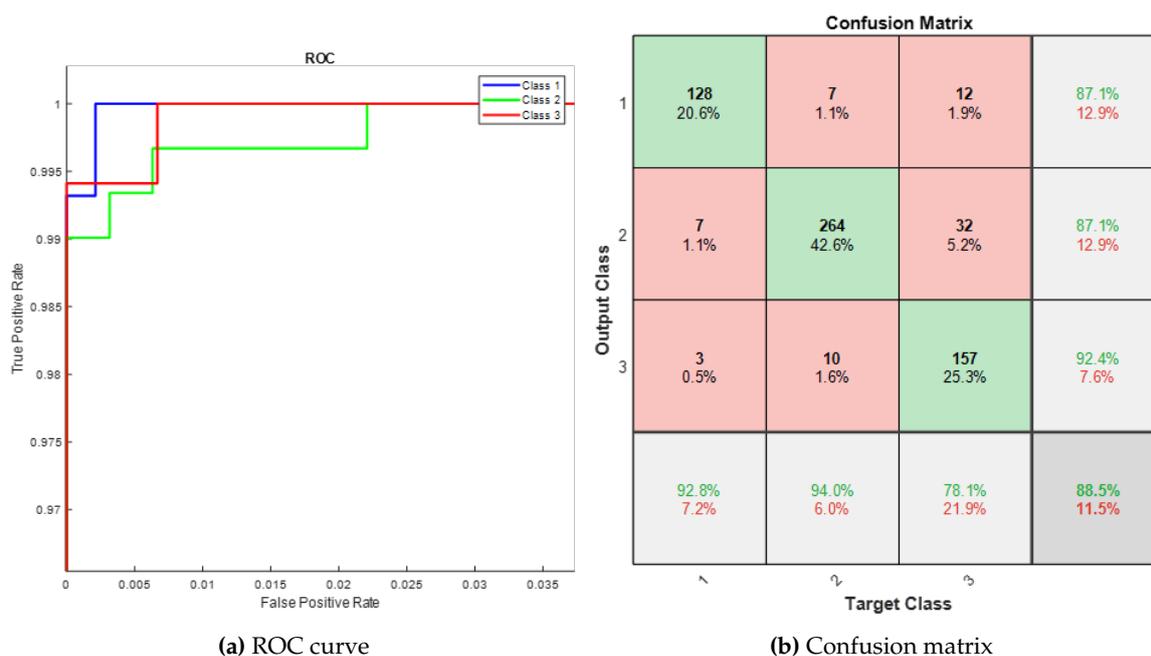


Figure 15. Co-pilot classifier performance.

It is important to note that most classification ‘confusion’ occurs between classes 2 and 3 (competent and expert pilots, respectively), which is not critical for the co-pilot application. The primary concern for the co-pilot application is the accurate identification of novice pilots (class 1), for which the classifier achieves an accuracy of 92.8%, demonstrating excellent performance.

The pilot skill level classification is performed in 0.02 seconds, theoretically allowing alerts to be generated at a rate of 50Hz. However, due to the 15-second latency required to accumulate a meaningful sequence length of flight parameters, such rapid alert generation is unnecessary. Therefore, a Kalman Filter was implemented to disregard misclassifications.

The output of the Kalman Filter is an auditory alert issued at a rate of 1 Hz to the drone operator when their flight behaviour is recognized by the classification system as corresponding to that of novice pilots. If the ‘bad’ flight behaviour is due to a lack of attention to the piloting task, the drone operator can take corrective measures. In cases where the pilot is indeed a novice still learning to operate a drone under an instructor’s supervision, the instructor can use the alerts from the virtual co-pilot to provide guidance on improving flight performance.

5.2. Automated Optimal Drone Trajectories for Target Observation

5.2.1. Methodology

To automatically generate optimal drone trajectories for target observation, we frame the problem as a constrained optimization task. Here, the objective function—such as the number of targets that need to be reached—is minimized while considering the constraints imposed by the drone’s flight dynamics. Consequently, it is essential to define the drone model and the application scenario, which in this case is a target observation mission.

In this study, we focus on rotorcraft drones, a practical choice since these types of unmanned aircraft are commonly used for short inspection or target observation tasks. Although executing complex dynamic flight behaviours with rotorcraft drones necessitates a sophisticated motion model and control architecture [83], a simplified motion model suffices for low-speed and relatively static observation applications relevant to this paper. Thus, we adopt a straightforward motion model [84] for generating potential locations for the drone to move to.

Additionally, we do not account for weather effects, such as wind, in our current model. While these external factors can be integrated into the system in future iterations, our primary aim is to validate the effectiveness of the proposed trajectory generation approach.

A pseudo-code representation of the general framework for generating drone trajectories is provided in Algorithm 1. The methodology is detailed line by line to elucidate the process.

- *Line 2:* As stated above, the algorithm starts from a simple drone motion model, which proposes a number of possible discrete locations where the drone can move to, taking into account the flight dynamics constraints.
In a first step, we perform a search over all possible new locations in order to assess which one is the best to move to. This means that a brute brute-force search is followed for searching for the optimal position. This is a quite simplistic approach, but we have opted for this option as the number of possible locations is not so enormous and it is therefore not required to incorporate some advanced optimization scheme.
- *Line 3:* In a second step, the safety of the proposed new drone location is assessed. This analysis considers in fact two different aspects:
 - The physical safety of the drone, which is in jeopardy if the drone comes too close to the ground. Therefore, a minimal distance from the ground will be imposed and proposed locations too close to the ground are disregarded.
 - The safety of the (stealth) observation operation, which is in jeopardy if the drone comes too close to the target, which means that the target (in a military context often an enemy)

could hear / perceive the drone and the stealthiness of the operation would thus be violated. Therefore, a minimal distance between the drone and the target will be imposed and proposed locations too close to the target will be disregarded.

- *Lines 4-6:* The different sub-criteria are assessed, following equations (1), (2) and (3).
- *Line 7:* The global objective video quality measure ϕ at the newly proposed location is calculated, following the equation (4).
- *Line 8:* The point with the highest video quality score ϕ is recorded.
- *Line 9-10:* At this point, an optimal point for the drone to move to has been selected (x_b). The viewpoint history memory θ_j and the velocity history memory \dot{x}_j are updated to include this new point.
- *Line 11:* The drone is moved to the new point x_b , in order to prepare for the next iteration.
- *Line 12:* The point x_b is appended to the drone trajectory profile.

Algorithm 1: Trajectory generation algorithm.

Input: drone position $\mathbf{x}_i = (x_i, y_i, z_i)$ target position $\mathbf{x}_t = (x_t, y_t, z_t)$

Output: Drone trajectory \mathbf{Y}

```

1 while not at end of the iteration do
2    $\mathbf{x}_n \leftarrow \text{CalculatePossibleNewPositions}(\mathbf{x}_i)$  forall proposed positions  $\mathbf{x}_n$  do
3     if EvaluateSafety( $\mathbf{x}_n$ ) then
4        $\phi_t \leftarrow \text{CalculatePixelsOnTarget}(\mathbf{x}_n, \mathbf{x}_t)$ 
5        $\phi_d \leftarrow \text{CalculateInnovation}(\mathbf{x}_n, \mathbf{x}_t, \theta_j)$ 
6        $\phi_s \leftarrow \text{CalculateSmoothness}(\mathbf{x}_n, \mathbf{x}_t, \dot{\mathbf{x}}_j)$ 
7        $\phi \leftarrow |(\phi_p, \phi_d, \phi_t)|$ 
8        $\mathbf{x}_b \leftarrow \text{RecordBestPoint}(\mathbf{x}_n, \phi)$ 
9    $\theta_j \leftarrow \text{UpdateDataInnovation}(\theta_j, \mathbf{x}_b)$ 
10   $\dot{\mathbf{x}}_j \leftarrow \text{UpdateTrajectorySmoothness}(\dot{\mathbf{x}}_j, \mathbf{x}_b)$ 
11   $\mathbf{x}_i \leftarrow \text{AdvanceDrone}(\mathbf{x}_b)$ 
12   $\mathbf{Y} \leftarrow \text{RecordDronePosition}(\mathbf{x}_b)$ 

```

5.2.2. Validation and Discussion

For the validation process, we initiated the drone from a random location and evaluated the optimal trajectories estimated by the algorithm.

An example of this analysis is illustrated in Figure 16. In the experiment depicted, the drone starts from a location which is significantly elevated above the target. The solution proposed by the proposed automatic trajectory generation methodology is shown in Figure 16e, where the target position is depicted by the large sphere on the bottom. Notably, the proposed solution involves a spiralling downward movement, ensuring comprehensive target perception from various angles. Upon reaching the safety distance from both the ground and the target, the movement pattern transitions to an outward-extending rectangular pattern. This movement pattern is both economical for the drone and ensures that the target is observed from increasingly oblique angles.

Figure 16a-16c depict the evolution of the sub-criteria ϕ_p , ϕ_d and ϕ_t throughout various stages of the drone's trajectory. As observed, the algorithm successfully achieves a relatively high number of pixels on target during the initial phase of the trajectory, while the drone spirals downward. In the subsequent phase, the number of pixels on target decreases as the drone moves further away to capture more oblique views.

The data innovation ϕ_d , illustrated in Figure 16b, exhibits a predominantly decreasing trend. This trend is attributed to the increasing difficulty in acquiring new information as the mission progresses.

Figure 16c demonstrates that the trajectory smoothness remains fairly constant for the majority of the trajectory. This consistency indicates the algorithm's effectiveness in selecting smooth trajec-

ries. However, near the end of the trajectory, peaks and valleys are observed, corresponding to the rectangular pattern where 90° turns alternate with straight paths.

By summing up the data innovation ϕ_d over time, a measure of scan completeness can be defined, as shown in Figure 16d. This metric provides an indication of the amount of new data collected per step of the trajectory. Across all conducted experiments, this scan completeness metric exhibits an asymptotic behaviour, as demonstrated in Figure 16d. This asymptotic trend is expected, as it becomes progressively more challenging to obtain new data over time. Consequently, this metric is invaluable for drone operators to evaluate in real-time whether it is worthwhile to continue the observation task or if it is more prudent to terminate the mission.

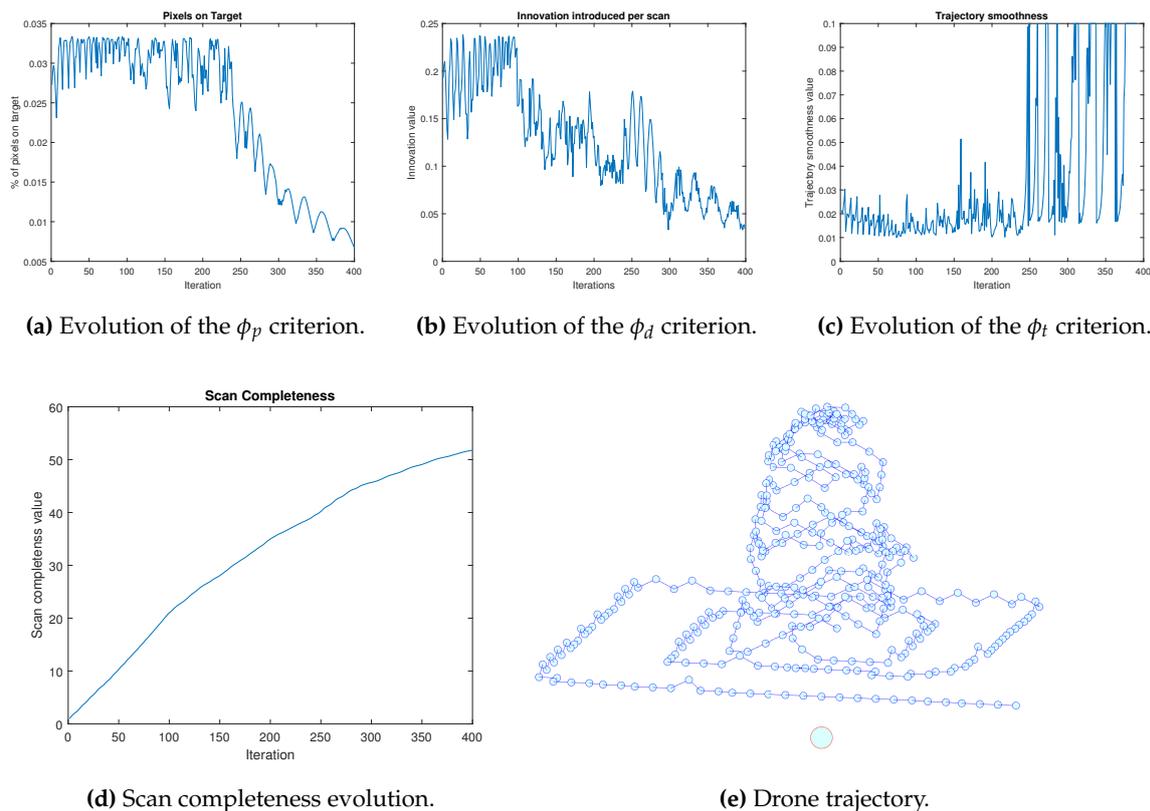


Figure 16. Drone trajectory generation validation.

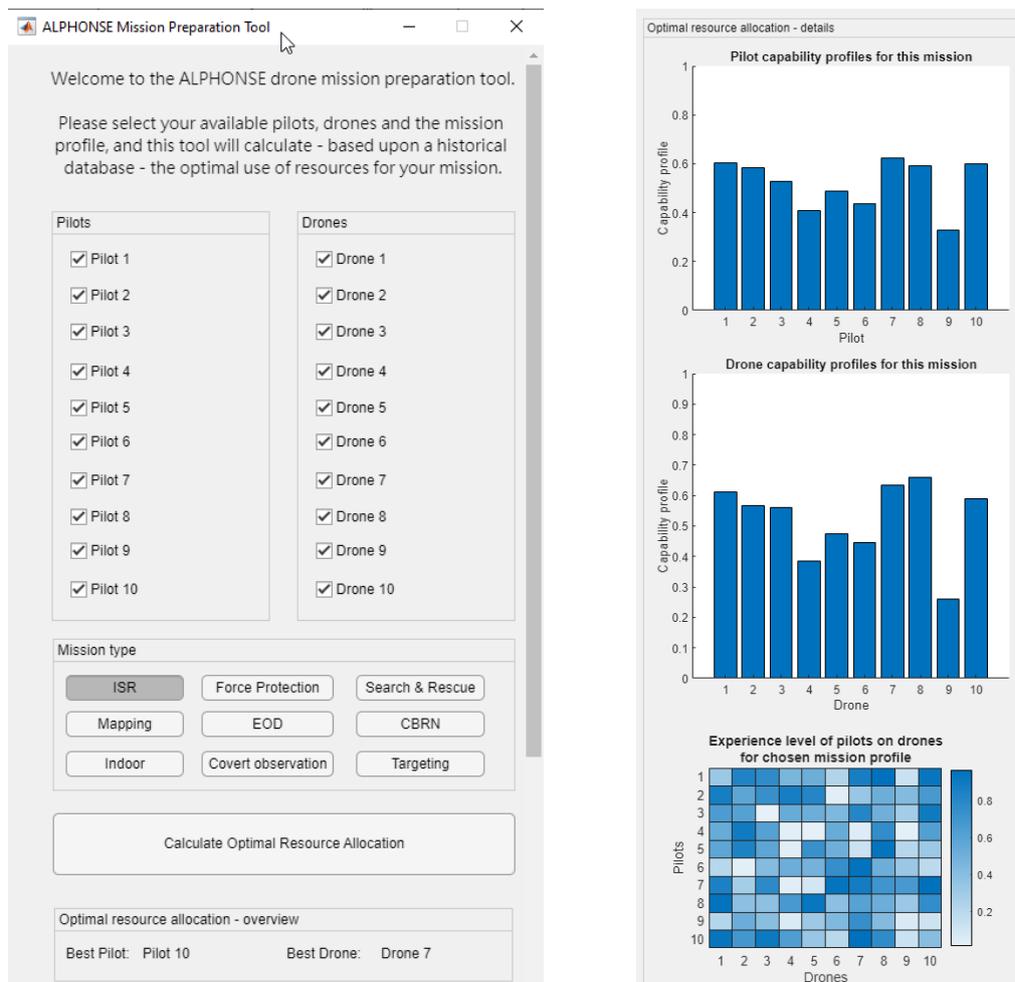
5.3. Drone Mission Planning Tool

Commanders in the field are often confronted with the situation that they have a number of human and technological assets (in this case: pilots and drones) at their disposal and they have to choose which assets to deploy for a specific mission. Using the quantitative performance data generated by the proposed assessment system, commanders can be offered a tool that enables an optimal resource allocation, based on statistical processing of quantitative data.

For this reason, an in-the-field drone mission preparation tool was developed, using the human performance data gathered using the pilot assessment tool. The tool, presented on the Figure 17a, allows to quickly select the pilots and drones that are available for a certain mission and to define the characteristics of the mission profile. At this moment, 9 pre-defined mission profiles are foreseen in the system: Intelligence Surveillance and Reconnaissance, Force protection, Search & rescue, Mapping, Explosive ordnance disposal, CBRN, Indoor, covert observation and targetting, as shown on Figure 17a. Based on historical performance data, the app will then propose the optimal resource allocation (pilot + drone) for the specified mission, as shown on the bottom of Figure 17a.

If the user wants more information, details on the resource allocation optimization can be consulted, as shown on the figure on the right. This can be insightful, as by analysing this information,

it can become clear that not in all cases the best overall drone pilot or best suited drone are chosen as an optimal combination. In the example below, it is e.g. evident that the chosen drone (number 7) performs overall less well than drone number 8. The final choice of the resource allocation tool for drone number 7 can be explained by the fact that it also takes into consideration the specific training & experience that each operator has with each drone (in practice: the number of flight hours, as shown on the bottom left of Figure 17b), as this is a key factor to be taken into consideration for the choice of the pilot and drone for any given mission.



(a) Interface of the drone mission preparation tool, enabling the operator to select the available drones and pilots and the type of mission. The system then calculates the optimal drone and pilot for the given mission and presents this information to the user.

(b) Details of the calculation of the optimal resource allocation. The optimizer uses stored capability profiles for drones and pilots for different missions and experience levels of the different pilots with all the drones to propose an optimal allocation.

Figure 17. Graphical user interface of the drone mission preparation tool.

5.4. Incremental Improvement of Drone Operator Training Procedures

5.4.1. Enabling Fine-Grained Pilot Accreditation

In manned aviation, there exist extremely strict procedures for pilot accreditation and aircraft airworthiness certification. For small unmanned aircraft, however, the rules are less tight and also less harmonized globally. In the European Union, a risk-based approach [85] is followed, where tighter rules are imposed (both for the pilot license as for the aircraft airworthiness assessment) with increasing risk associated to the drone operation to be performed. A crucial point is thus to assess the

risk to a drone operation, which is dependent of the scenarios that are going to be performed and that are written down in the operational handbook. Therefore, a set of standard scenarios are defined and in order to get a permission to fly, the performance of drone pilots and drones for a specific scenario needs to be assessed. This concept of operation for the accreditation has an important pitfall that our work tries to address: the drone pilot accreditation process happens once, once a year or once every few years. However, we know that a varying physiological state of the pilot on the date of the flight may impact the performance drastically. Using our human performance model, we can predict – given a certain physiological input state – what would be the flight performance of the human operator. As such, a much more fine-grained case-based accreditation is possible, which is specifically useful for stressful operations, such as is often the case in the security sector.

5.4.2. Enabling Iterative Improvement of Training Procedures

The described system represents a significant advancement in the domain of drone pilot training, offering a dynamic and iterative methodology for performance enhancement. This approach can be effectively employed by training agencies to refine and elevate their training procedures continuously. Central to this system is the ability to capture and analyse extensive performance data within a virtual test environment, allowing for a comprehensive assessment of pilot skills under various conditions.

Training agencies can leverage the system's detailed data collection capabilities to conduct an in-depth analysis of pilot performance, identifying patterns and pinpointing specific areas where pilots may struggle. This data-driven insight enables the development of tailored training programs that address individual weaknesses and promote overall proficiency. Furthermore, the integration of a human performance model provides a nuanced understanding of how factors such as stress and cognitive load impact pilot effectiveness. By incorporating these insights, training programs can be adjusted to mitigate these factors, thereby enhancing pilot performance.

The real-time feedback provided by the AI co-pilot is another critical asset, offering immediate guidance and correction during training sessions. This feature ensures that pilots can rectify mistakes as they occur, fostering the development of correct practices and preventing the reinforcement of poor behaviours. The continuous feedback loop created by this system is instrumental in promoting iterative improvement, as it allows for the ongoing refinement of training techniques based on real-time performance data.

Moreover, the system's mission planning tool plays a crucial role in optimizing pilot assignment and training scenarios. By simulating real-world mission planning and assessing pilot suitability based on performance scores and stress models, the tool prepares pilots for actual operations and enhances their decision-making skills. This aspect of the system ensures that training remains relevant and aligned with operational demands, providing pilots with a realistic and practical training experience.

The iterative nature of this system is further reinforced by the continuous incorporation of quantitative input from training sessions into the overall training curriculum. As pilots engage with the training modules, their performance data is systematically analysed and used to update and improve the training content. This process ensures that the training programs evolve in response to the latest performance insights, maintaining their effectiveness and relevance over time.

In essence, the described system enables training agencies to adopt a holistic and adaptive approach to drone pilot training. By utilizing comprehensive data analysis, real-time feedback, and iterative curriculum updates, training programs can be continuously refined to produce highly skilled and proficient drone pilots. This approach not only enhances individual pilot performance but also contributes to the overall advancement of training methodologies within the field.

5.4.3. Enabling Fine-Grained Pilot Performance Follow-Up

The presented approach allows training agencies to monitor pilot progress with precision, provide targeted interventions, and continuously refine training programs to enhance pilot performance comprehensively. The simulator system facilitates fine-grained pilot performance follow-up through

its detailed data collection and analysis capabilities. By capturing extensive performance metrics within a sophisticated virtual test environment, the system provides a comprehensive and nuanced understanding of each pilot's abilities and areas for improvement.

Training agencies can utilize this detailed data to monitor pilot performance over time, allowing for precise tracking of progress and identification of trends. This longitudinal data collection enables a granular analysis of how specific skills develop and how different factors, such as training interventions or operational stressors, impact performance. By continuously analysing this data, agencies can provide targeted feedback and tailor training programs to address individual pilot needs.

Moreover, the iterative nature of the system's data integration ensures that performance follow-up is a continuous and evolving process. As pilots engage in training, their performance data is regularly fed back into the system, allowing for ongoing updates and refinements to the training curriculum. This iterative feedback loop ensures that training programs remain adaptive and responsive to the latest performance insights, providing pilots with continuously optimized training experiences.

6. Conclusions

6.1. Discussion on the Proposed Contributions

In this paper, we propose a novel drone operator performance assessment tool that leverages realistic environments and operational conditions to measure operator performance both qualitatively and quantitatively. This tool aims to optimize training curricula to ensure maximum safety. Realism in simulation systems is crucial for achieving desired training outcomes. Our proposed framework incorporates realistic operational conditions, such as wind and weather effects, which are critical for assessing pilot performance in complex scenarios. This approach is particularly relevant for pilots in challenging conditions, such as those in military or emergency services, who find simplistic training scenarios inadequate.

The development of our performance assessment tool includes a sophisticated simulator built on the Microsoft AirSim engine and the Unreal Engine. This simulator tracks over 65 flight parameters and assesses video quality, providing a highly realistic environment for complex operations. It also suggests optimal flight trajectories for target observation, thereby assisting operators during missions. Post-flight, an overall performance score is calculated based on mission objectives. This data has been used to train an AI-based classifier to recognize 'good' and 'bad' flight behaviours in real-time, which aids in the development of a virtual AI co-pilot for immediate feedback.

Furthermore, our research has led to the development of a mission planning tool that assigns the best pilots for specific missions based on performance scores. This tool helps identify and rank human factors impacting flight performance. It was also the goal to link stress factors to performance outcomes, but as shown by Figure 12, we were not able to deduct such a relationship, such that we couldn't include stress in the model.

Our study makes several novel contributions to the field: a drone simulator that tracks operator performance under realistic conditions, a method for quantitative evaluation of drone-based video quality, and a comprehensive human performance modelling methodology that links human factors to operator performance. Additionally, the introduction of an AI co-pilot for real-time flight performance guidance and a flight assistant tool for optimal trajectory generation represents significant advancements in drone pilot training technologies. Furthermore, by training a range of pilots in this simulator environment, these innovations enable the development of several applications, including an iterative training improvement methodology based on quantitative input. These developments shown in this research work are developed to meet the needs of the Belgian Special Forces Group, focusing on scenarios such as ISR operations. However, the presented approach is totally generic and open, and therefore highly adaptable to various other use cases.

6.2. Future Work

While the presented drone operator performance assessment tool represents a significant advancement in the field of drone pilot training, it remains only an *Alpha* version and can benefit from multiple potential future improvements to further enhance its effectiveness and applicability.

One promising direction is the continuous refinement and expansion of the simulation scenarios. While the current simulator effectively incorporates an ISR scenario in a mountainous environment, future iterations should include an even broader range of scenarios and environmental variables and emergency situations. This would provide a more comprehensive training experience, better preparing pilots for the wide variety of conditions they may encounter in real-world operations. Additionally, expanding the range of simulated scenarios to include more specific mission types, such as search and rescue operations, civil security, or industrial inspections, could tailor the training experience to the needs of various industries. This would help bridge the gap between simulation and real-world operations, leading to more effective training outcomes.

Another potential area of future work is the integration of more advanced feedback systems. While the current system relies on intake and outtake questionnaires, future versions could incorporate exteroceptive sensing to gather quantitative stress and task load data from the pilots during the flight operation, which would enable correlating specific flight operations with stress levels. Potentially, this could also alleviate the problem that we failed to identify (in Figure 12) a correlation between the stress level and the pilot flight performance.

What is certainly required to extend the system applicability is to incorporate more drone pilots in the testing process. Moreover, longitudinal studies tracking the progress of pilots over extended periods could provide valuable insights into the long-term effectiveness of the training program and identify areas for further refinement. Additionally, collaborative research with psychological and cognitive scientists could help to better understand the human factors influencing drone pilot performance and how best to address them in training programs.

Another avenue for future work is the development of better metrics for assessing drone pilot performance, as the current metrics are basically a first guess by the authors. Establishing a set of more universally accepted performance indicators would facilitate more consistent and objective evaluations across different training programs and operational contexts. This could involve collaboration with regulatory bodies and industry stakeholders to ensure that the metrics are comprehensive and applicable to a wide range of drone operations.

Lastly, expanding the applicability of the simulator to a broader audience, including civilian and recreational drone pilots, could have significant implications for drone safety and regulation. Developing user-friendly versions of the simulator with customizable training modules could help democratize access to high-quality drone training, ultimately contributing to safer skies for all drone operators.

Author Contributions: Conceptualization, D. Doroftei, G. De Cubber and S. Lo Bue; methodology, D. Doroftei and G. De Cubber; software, D. Doroftei and G. De Cubber; validation, D. Doroftei; formal analysis, D. Doroftei; investigation, D. Doroftei; resources, D. Doroftei; data curation, D. Doroftei; writing—original draft preparation, D. Doroftei; writing—review and editing, G. De Cubber, S. Lo Bue and H. De Smet; visualization, D. Doroftei; supervision, H. De Smet; project administration, H. De Smet; funding acquisition, H. De Smet. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Belgian Royal Higher Institute for Defense in the framework of the research study HFM19/05 (ALPHONSE).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The datasets presented in this article are not readily available because they include operational performance data of Belgian Defence military operatives. Publicly releasing this data would thus constitute a security concern. Requests to access the datasets should be directed to the corresponding author.

Acknowledgments: We extend our gratitude to the VIAS Institute for their support throughout this study, both by providing essential equipment and offering invaluable advice and guidance during the project's execution.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ADAM	Adaptive Moment Estimation
AI	Artificial Intelligence
APAS	Advanced Pilot Assistance Systems
ATM	Aerial Traffic Management
BVLOS	Beyond Visual Line Of Sight
CBRN	Chemical, Biological, Radiological and Nuclear
CSV	Comma-separated values
DEG	Dynamic Environment Generator
DJI	Da-Jiang Innovations
DRL	Drone Racing League
GB	Gigabyte
GPS	Global Positioning System
GUI	Graphical User Interface
HITL	Hardware In The Loop
HMI	Human Machine Interface
ISR	Intelligence, Surveillance, Reconnaissance
LiDAR	Light Detection and Ranging
MDPI	Multidisciplinary Digital Publishing Institute
MTE	Mission Task Element
NIST	National Institute of Standards and Technology
PSD	Power Spectral Density
PSNR	Peak Signal-to-Noise Ratio
R-CNN	Region Convolutional Neural Network
RC	Remote Control
RF	Radio Frequency
RMSE	Root Mean Square Error
ROC	Receiver Operating Characteristic
SITL	Software In The Loop
TCP	Transmission Control Protocol
TLX	Task Load Index
UAS	Unmanned Aircraft System
UAVs	Unmanned Aerial Vehicle
UDP	User Datagram Protocol
UTM	Unmanned Traffic Management
YOLO	You Only Look Once)

References

1. Chow, E.; Cuadra, A.; Whitlock, C. Hazard Above: Drone Crash Database - Fallen from the skies. The Washington Post, 2016.
2. Buric, M.; De Cubber, G. Counter Remotely Piloted Aircraft Systems. *MTA Review* **2017**, *27*, No. 1, Military Technical Academy Publishing House.
3. Shively, J. Human Performance Issues in Remotely Piloted Aircraft Systems. In Proceedings of the ICAO Conference on Remotely piloted or piloted: sharing one aerospace system, Montreal, Canada, 23-25 March 2015.
4. Lee, J.D.; Wickens, C.D.; Liu, Y.; Boyle, L.N. Designing for people: An introduction to human factors engineering. Charleston:CreateSpace, 2017.
5. Fletcher, G. Pilot training review – interim report: literature review, Published by the British Civil Aviation Authority, 2017.

6. Socha, V.; Socha, L.; Szabo, S.; Hana, K.; Gazda, J.; Kimlickova, M.; Vajdova, I.; Madoran, A.; Hanakova, L.; Nemec, V. Training of pilots using flight simulator and its impact on piloting precision. In *Transport Means. Juodkrante: Kansas University of Technology*. 2016, 374–379.
7. Rostáš, J.; Kováčiková, M.; Kandra, B. Use of a simulator for practical training of pilots of unmanned aerial vehicles in the Slovak Republic. In Proceedings of the 19th IEEE International Conference on Emerging eLearning Technologies and Applications (ICETA), Košice, Slovakia, 11-12 November 2021, pp. 313-319, IEEE.
8. Shah, S.; Dey, D.; Lovett, C.; Kapoor, A. AirSim: High-Fidelity Visual and Physical Simulation for Autonomous Vehicles. In *Field and Service Robotics. Springer Proceedings in Advanced Robotics, vol 5*; Hutter, M.; Siegart, R., Eds.; Springer, Cham; 2017.
9. Lee, N. Unreal Engine, A 3D Game Engine. In: *Encyclopedia of Computer Graphics and Games.*; Lee, N. (eds); Springer, Cham; 2023.
10. Mairaj, A.; Baba, A.I.; Javaid, A.Y. Application specific drone simulators: Recent advances and challenges, *Simulation Modelling Practice and Theory* 2019, 94.
11. DJI Flight Simulator. Available online: <https://www.dji.com/be/downloads/products/simulator> (accessed on 16 June 2024).
12. The Drone Racing League Simulator. Available online: https://store.steampowered.com/app/641780/The_Drone_Racing_League_Simulator/ (accessed on 16 June 2024).
13. Zephyr. Available online: <https://zephyr-sim.com/> (accessed on 16 June 2024).
14. droneSim Pro. Available online: <https://www.dronesimpro.com/> (accessed on 16 June 2024).
15. RealFlight. Available online: <https://www.realflight.com/product/realflight-9.5s-rc-flight-sim-with-interlink-controller/RFL1200S.html> (accessed on 16 June 2024).
16. Cotting, M. An initial study to categorize unmanned aerial vehicles for flying qualities evaluation. In Proceedings of the 47th AIAA Aerospace Sciences Meeting including The New Horizons Forum and Aerospace Exposition, 2009..
17. Holmberg, J.; Leonard, J.; King, D.; Cotting, M. Flying qualities specifications and design standards for unmanned air vehicles. In Proceedings of the AIAA Atmospheric Flight Mechanics Conference and Exhibit, 2008.
18. Hall, C.; Southwell, J. Equivalent Safe Response Model for Evaluating the Closed Loop Handling Characteristics of UAS to Contribute to the Safe Integration of UAS into the National Airspace System. In Proceedings of the 11th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference, including the AIAA Balloon Systems Conference, 2011.
19. Schulze, P. C.; Miller, J.; Klyde, D. H.; Regan, C. D.; and Alexandrov, N. System Identification of a Small UAS in Support of Handling Qualities Evaluations. In Proceedings of the AIAA Scitech 2019 Forum, 2019.
20. Abdulrahim, M.; Bates, T.; Nilson, T.; Bloch, J.; Nethery, D.; and Smith, T. Defining Flight Envelope Requirements and Handling Qualities Criteria for First-Person-View Quadrotor Racing. In Proceedings of the AIAA Scitech 2019 Forum, 2019.
21. Greene, K. M.; Kunz, D. L.; Cotting, M. C. Toward a Flying Qualities Standard for Unmanned Aircraft. In Proceedings of the AIAA Atmospheric Flight Mechanics Conference, 2014.
22. Klyde, D. H.; Schulze, P. C.; Mitchell, D.; and Alexandrov, N.. Development of a Process to Define Unmanned Aircraft Systems Handling Qualities. In Proceedings of the AIAA Atmospheric Flight Mechanics Conference, 2018.
23. Sanders, F. C.; Tischler, M.; Berger, T.; Berrios, M. G.; and Gong, A. System Identification and Multi-Objective Longitudinal Control Law Design for a Small Fixed-Wing UAV. In Proceedings of the AIAA Atmospheric Flight Mechanics Conference, 2018.
24. Abdulrahim, M.; Dee, J.; Thomas, G.; and Qualls, G. Handling Qualities and Performance Metrics for First-Person-View Racing Quadrotors. In Proceedings of the AIAA Atmospheric Flight Mechanics Conference, 2018.
25. Herrington, S.M.; Hasan Zahed, M.J.; Fields, T. Pilot Training and Task Based Performance Evaluation of an Unmanned Aerial Vehicle. In Proceedings of the AIAA Scitech 2021 Forum, 11–21 January 2021.
26. Ververs, P.M.; Wickens, C.D., Head up displays: effect of clutter, display intensity and display location of pilot performance. *Int. J. Aviat. Psychol.* 1998, 8(4), 377–403

27. Smith, J. K.; Caldwell, J. A. Methodology for evaluating the simulator flight performance of pilots. Brooks City Base, TX: Air Force Research Laboratory. Report No, AFRL-HE-BR-TR-2004-0118, 2004
28. Hanson, C.; Schaefer, J., Burken, J. J.; Larson, D.; Johnson, M. Complexity and pilot workload metrics for the evaluation of adaptive flight controls on a full scale piloted aircraft. NASA Dryden Flight Research Center; Edwards, CA United States. Document ID. 20140005730, 2014.
29. Field, E. J.; Giese, S. E. D. Appraisal of several pilot control activity measures. In Proceedings of the AIAA Atmospheric Flight mechanics conference and exhibit, San Francisco, California, 15–18 August 2005.
30. Zahed, M.J.H.; Fields, T. Evaluation of pilot and quadcopter performance from open-loop mission-oriented flight testing. *Journal of Aerospace Engineering* **2021**, *235(13)*, pp 1817–1830.
31. Hebbar, P.A.; Pashilkar, A.A. Pilot performance evaluation of simulated flight approach and landing manoeuvres using quantitative assessment tools. *Sādhanā Academy Proceedings in Engineering Sciences* **2017**, *42*, pp 405–415.
32. Jacoff, A.; Mattson, P. Measuring and comparing small unmanned aircraft system capabilities and remote pilot proficiency. National Institute of Standards and Technology, 2020.
33. Hoßfeld, T.; Keimel, C.; Hirth, M.; Gardlo, B.; Habigt, J.; Diepold, K.; Tran-Gia, P. Best practices for qoe crowdtesting: Qoe assessment with crowdsourcing. *IEEE Transactions on Multimedia* **2014**, *16*, pp. 541–558.
34. Takahashi, A.; Hands, D.; Barriac, V. Standardization activities in the ITU for a QoE assessment of IPTV. *IEEE Communications Magazine* **2008**, *46*, pp. 78–84.
35. Winkler, S.; Mohandas, P. The evolution of video quality measurement: From PSNR to hybrid metrics. *IEEE Transactions on Broadcasting* **2008**, *54*, pp. 660–668.
36. Hulens, D.; Goedeme, T. Autonomous flying cameraman with embedded person detection and tracking while applying cinematographic rules. In Proceedings of the 14th Conference on computer and robot vision (CRV2017), Edmonton, Canada, 16-19 May 2017, pp. 56–63.
37. Doroftei, D.; De Cubber, G.; De Smet, H. A quantitative measure for the evaluation of drone-based video quality on a target. In Proceedings of the Eighteenth International Conference on Autonomic and Autonomous Systems (ICAS), Venice, Italy, 22-26 May 2022.
38. Deutsch, S. UAV Operator Human Performance Models. BBN Report 8460, 2006.
39. Bertuccelli, L.F.; Beckers, N.W.M.; Cummings, M.L. Developing operator models for UAV search scheduling. In Proceedings of the of AIAA Guidance, Navigation, and Control Conference, Toronto, Canada, 2010.
40. Wu, Y.; Huang, Z.; Li, Y.; Wang, Z. Modeling Multioperator Multi-UAV Operator Attention Allocation Problem Based on Maximizing the Global Reward. *IEEE Mathematical Problems in Engineering* **2016**.
41. Hart, S. G.; Staveland, L. E. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research" (PDF). In *Human Mental Workload. Advances in Psychology*, Hancock, P. A.; Meshkati, N. (eds.). *52*, pp. 139–183.
42. Andrews, Jinan M. Human Performance Modeling: Analysis of the Effects of Manned-Unmanned Teaming on Pilot Workload and Mission Performance. Air Force Institute of Technology Theses and Dissertations, 2020.
43. Wright, J.L.; Lee, J.; Schreck, J.A. Human-autonomy teaming with learning capable agents: performance and workload outcomes. In Proceedings of the International Conference on Applied Human Factors and Ergonomics, 2021.
44. Doroftei, D.; De Cubber, G.; De Smet, H. Human factors assessment for drone operations: towards a virtual drone co-pilot. In Proceedings of the AHFE International Conference on Human Factors in Robots, Drones and Unmanned Systems, 2023.
45. Sakib, M.N.; Chaspari, T.; Ahn, C.; Behzadan, A. An experimental study of wearable technology and immersive virtual reality for drone operator training. In Proceedings of the 27th Int. Workshop on Intelligent Computing in Engineering, 2020, pp. 154–163.
46. Sakib, M.N. Wearable technology to assess the effectiveness of virtual reality training for drone operators (Doctoral dissertation), 2019.
47. Doroftei, D.; De Cubber, G.; De Smet, H. Reducing drone incidents by incorporating human factors in the drone and drone pilot accreditation process. In Proceedings of the AHFE 2020 Virtual Conference on Human Factors in Robots, Drones and Unmanned Systems, 16-20 July 2020, USA, 2020, pp. 71–77.
48. Gupta, S. G.; Ghonge, M.; Jawandhiya, P. M. Review of unmanned aircraft system (UAS). *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, **2013**, *2*.

49. Hussein, M.; Nouacer, R.; Corradi, F.; Ouhammou, Y.; Villar, E.; Tieri, C.; Castiñeira, R. Key technologies for safe and autonomous drones. *Microprocessors and Microsystems*, **2021**, *87*.
50. Chandran, N. K.; Sultan, M. T. H.; Łukaszewicz, A.; Shahar, F. S.; Holovatyy, A.; Giernacki, W. Review on Type of Sensors and Detection Method of Anti-Collision System of Unmanned Aerial Vehicle. *Sensors*, **2023**, *23*(15).
51. Gupta, A.; Fernando, X. Simultaneous localization and mapping (slam) and data fusion in unmanned aerial vehicles: Recent advances and challenges. *Drones*, **2022**, *6*(4).
52. Castro, G. G. D.; Berger, G. S.; Cantieri, A.; Teixeira, M.; Lima, J.; Pereira, A. I.; Pinto, M. F. Adaptive path planning for fusing rapidly exploring random trees and deep reinforcement learning in an agriculture dynamic environment UAVs. *Agriculture*, **2023**, *13*(2).
53. Telli, K.; Kraa, O.; Himeur, Y.; Ouamane, A.; Boumehraz, M.; Atalla, S.; Mansoor, W. A comprehensive review of recent research trends on unmanned aerial vehicles (uavs). *Systems*, **2023**, *11*(8).
54. Azar, A. T.; Koubaa, A.; Ali Mohamed, N.; Ibrahim, H. A.; Ibrahim, Z. F.; Kazim, M.; Casalino, G. Drone deep reinforcement learning: A review. *Electronics*, **2021**, *10*(9).
55. Jawaharlalnehru, A.; Sambandham, T.; Sekar, V.; Ravikumar, D.; Loganathan, V.; Kannadasan, R.; Alzamil, Z. S. Target object detection from Unmanned Aerial Vehicle (UAV) images based on improved YOLO algorithm. *Electronics*, **2022**, *11*(15).
56. McConville, A.; Bose, L.; Clarke, R.; Mayol-Cuevas, W.; Chen, J.; Greatwood, C.; Richardson, T. Visual odometry using pixel processor arrays for unmanned aerial systems in gps denied environments. *Frontiers in Robotics and AI*, **2020**, *7*.
57. van de Merwe, K.; Mallam, S.; Nazir, S. Agent transparency, situation awareness, mental workload, and operator performance: A systematic literature review. *Human Factors*, **2024**, *66*(1), pp. 180–208.
58. Woodward, J.; Ruiz, J. Analytic review of using augmented reality for situational awareness. *IEEE Transactions on Visualization and Computer Graphics*, **2022**, *29*(4), pp. 2166–2183.
59. Van Baelen, D.; Ellerbroek, J.; Van Paassen, M. M.; Mulder, M. Design of a haptic feedback system for flight envelope protection. *Journal of Guidance, Control, and Dynamics*, **2020**, *43*(4), pp 700–714.
60. Nguyen, T-T.; Crismer, A.; De Cubber, G.; Janssens, B.; Bruyninckx, H. Landing UAV on Moving Surface Vehicle: Visual Tracking and Motion Prediction of Landing Deck. In Proceedings of the IEEE/SICE International Symposium on System Integration (SII), 2024.
61. Dutrannois, T.; Nguyen, T-T.; Hamesse, C.; De Cubber, G.; Janssens, B. Visual SLAM for Autonomous Drone Landing on a Maritime Platform. In Proceedings of the International Symposium for Measurement and Control in Robotics (ISMCR) – A Topical Event of Technical Committee on Measurement and Control of Robotics (TC17), International Measurement Confederation (IMEKO), 2022.
62. Papyan, N.; Kulhandjian, M.; Kulhandjian, H.; Aslanyan, L. AI-Based Drone Assisted Human Rescue in Disaster Environments: Challenges and Opportunities. *Pattern Recognition and Image Analysis*, **2024**, *34*(1), pp. 169–186.
63. Weber, U.; Attinger, S.; Baschek, B.; Boike, J.; Borchardt, D.; Brix, H.; Brüggemann, N.; Bussmann, I.; Dietrich, P.; Fischer, P.; Greinert, J. MOSES: A novel observation system to monitor dynamic events across Earth compartments. *Bulletin of the American Meteorological Society*, **2022**, *103*(2), pp.339–348.
64. Ramos, M. A.; Sankaran, K.; Guarro, S.; Mosleh, A.; Ramezani, R.; Arjounilla, A. The need for and conceptual design of an AI model-based Integrated Flight Advisory System. *Journal of Risk and Reliability*, **2023**, *237*(2), pp. 485–507.
65. Doroftei, D.; De Cubber, G.; De Smet, H. Assessing Human Factors for Drone Operations in a Simulation Environment. In Proceedings of the Human Factors in Robots, Drones and Unmanned Systems – AHFE (2022) International Conference, New York, USA, 2022.
66. Doroftei, D.; De Smet, H. Evaluating Human Factors for Drone Operations using Simulations and Standardized Tests. In Proceedings of the 10th International Conference on Applied Human Factors and Ergonomics (AHFE 2019), Washington DC, USA, 2019.
67. De Cubber, G.; Berrabah, S. A.; Sahli, H. Color-based visual servoing under varying illumination conditions. *Robotics and Autonomous Systems*, **2004**, *47*(4), pp. 225–249.
68. De Cubber, G.; Marton, G. Human victim detection. in Proceedings of the Third International Workshop on Robotics for risky interventions and Environmental Surveillance-Maintenance, RISE, 2009.

69. Marques, J. S.; Bernardino, A.; Cruz, G.; Bento, M. An algorithm for the detection of vessels in aerial images. in Proceedings of the 11th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), pp. 295–300, Aug 2014.
70. Johnson, J. Analysis of image forming systems. in Proceedings Image Intensifier Symposium, p. 244–273, 1958.
71. Meier, L.; Honegger, D.; Pollefeys, M. PX4: A node-based multithreaded open source robotics framework for deeply embedded platforms. In 2015 IEEE international conference on robotics and automation (ICRA), 2015, pp. 6235–6240.
72. Doroftei, D.; De Vleeschauwer, T.; Lo Bue, S.; Dewyn, M.; Vanderstraeten, F.; De Cubber, G. Human-Agent Trust Evaluation in a Digital Twin Context. In 30th IEEE International Conference on Robot Human Interactive Communication (RO-MAN), Vancouver, BC, Canada, 2021, pp. 203–207.
73. Doroftei, D.; De Cubber, G.; Wagemans, R.; Matos, A.; Silva, E.; Lobo, V.; Guerreiro Cardoso, K. C.; Govindaraj, S.; Gancet, J.; Serrano, D. User-centered design. In Search and rescue robotics. From theory to practice. De Cubber, G.; Doroftei, D.; Eds. IntechOpen, London, 2017, p. 19–36.
74. Holmes, T.H.; Rahe, T.H. The Social Readjustment Rating Scale. *Journal of Psychosomatic Research*, **1967**, *11*(213).
75. Koubâa, A.; Allouch, A.; Alajlan, M.; Javed, Y.; Belghith, A.; Khalgui, M. Micro Air Vehicle Link (MAVlink) in a Nutshell: A Survey. in *IEEE Access*, **2019**, *7*, pp. 87658–87680.
76. mavp2p. Available online: <https://github.com/bluenvirion/mavp2p> (accessed on 03 July 2024).
77. Garinther, G.R.; Kalb, I.J.T.; Hodge, D.C.; Price, G.R. Proposed aural non-detectability limits for army materiel, U.S. Army Human Engineering Laboratory, 1985
78. Doroftei, D.; De Cubber, G. Using a qualitative and quantitative validation methodology to evaluate a drone detection system, *ACTA IMEKO*, **2019** *8*(4), pp. 20–27.
79. De Cubber, G.; Shalom, R.; Coluccia, A.; Borcan, O.; Chamrád, R.; Radulescu, T.; Izquierdo, E.; Gagov, Z. The SafeShore system for the detection of threat agents in a maritime border environment. In IARP Workshop on Risky Interventions and Environmental Surveillance (Vol. 2), 2017.
80. Ramirez-Atencia, C.; Camacho, D. Extending QGroundControl for automated mission planning of UAVs. *Sensors*, **2018**, *18*(7).
81. Bishop, C. M. In *Pattern Recognition and Machine Learning*. Springer, New York, NY, USA, 2006.
82. Kingma, D.; Ba, J. Adam: A method for stochastic optimization. 3rd International Conference on Learning Representations, 2015.
83. Alexis, K.; Nikolakopoulos, G.; Tzes, A. Model predictive quadrotor control: attitude, altitude and position experimental studies. *IET Control Theory & Applications*, **2012**, *6*(12), 1812–1827.
84. Szolc, H.; Kryjak, T. Hardware-in-the-loop simulation of a UAV autonomous landing algorithm implemented in SoC FPGA. In 2022 Signal Processing: Algorithms, Architectures, Arrangements, and Applications, pp. 135–140.
85. European Commission. Commission Implementing Regulation (EU) 2019/947 of 24 May 2019 on the rules and procedures for the operation of unmanned aircraft.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.