Human Performance Envelope Model Study Using Pilot's Measured Parameters

Alina-Ioana CHIRA*,[1], Anamaria DUMITRESCU[1], Catalin Sever MOISOIU[1], Cristian-Alexandru TANASE[1]

*Corresponding author
[1]INCAS – National Institute for Aerospace Research "Elie Carafoli", Blvd Iuliu Maniu 220, Bucharest 061126, Romania,
chira.alina@incas.ro, dumitrescu.anamaria@incas.ro, moisoiu.catalin@incas.ro, tanase.alexandru@incas.ro

DOI: 10.13111/2066-8201.2020.12.4.5

Received: 18 October 2020/ Accepted: 05 November 2020/ Published: December 2020

International Conference of Aerospace Sciences "AEROSPATIAL 2020", Virtual Conference 15-16 October 2020, Bucharest, Romania,
Section 5 - Systems, Subsystems and Control in Aeronautics

Abstract: *Taking into consideration that nowadays the aerospace industry focuses a lot on safety, more durable and stable systems are developed. While the system itself is safer, there is another element that can have a high impact on the overall safety of a flight, namely the human factor. Pilot physiological parameters were measured during a full flight in a fixed cockpit environment using application-specific equipment. The recorded or calculated parameters are used to compute a performance envelope model with the scope of determining the degradation of the pilot's condition during different flight phases or events. Several standardized tests were realized/performed on subjects who were given flight instructions before the test, without knowing beforehand the scenario and events that will occur/take place. This study helps in identifying the limits of pilots in different flight scenarios and the impact on their presumed performance.*

Key Words: *human factors, safety, pilots, flight simulator, eye tracking, heatmap, ECG, HPE*

1. INTRODUCTION

As in any other field, accidents happen, especially aviation accidents that can very easily turn into real catastrophes, simply because of the conditions in which these take place.

Approximately 75 - 80 percent of airplane accidents are due to human error, as shown in Fig. 1 [1].

According to Boeing's Statistical Summary of Commercial Jet Airplane Accidents, aeronautical the 2007-2016 aviation accidents with human factors, technical failure and weather conditions happened in the final phases/stages of the flights, 48% being fatal accidents.

These phases rely mostly on the human pilots without the possibility of always having an autopilot mode [2].

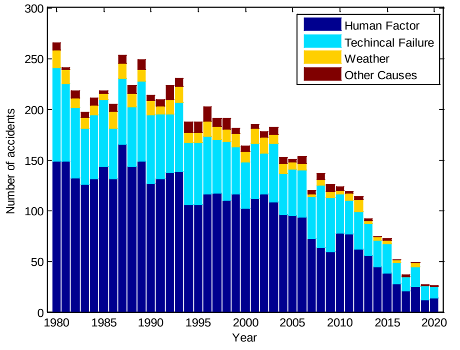


Fig. 1 - The 1980-2020 aviation accidents with human factors, technical failures, weather and other causes
Data retrieved from the Bureau of Aircraft Accidents Archives [1]

In this article, we analyzed the behavior of the pilots during a full flight with the duration of about one hour. Along some of the flights, errors were introduced without prior notice to the pilots. The scope of these errors was to generate changes in physiological parameters that were recorded using application-specific equipment.

Using the recorded parameters, several factors that contribute to performance were calculated. The main objective of the design experiment was to develop a series of scenarios that utilize the same flight task, but which could require the pilot several different levels of involvement.

To achieve these scenarios, a full flight task from the International Airport Henri Coanda Bucharest (LROP) to Timisoara Traian Vuia Airport (LRTR) was chosen.

In order to quantify the recorded physiological parameters, a human performance envelope model was computed. Our Human Performance Envelope (HPE) model is composed of Workload (WL), Stress, Situational Awareness (SA) and Attention.

Human Performance Envelope represents a graphical representation of a group of interdependent factors that add to a performance calculation. We identified several components of the physiological measurement that contribute to each factor in HPE. In Table 1, we noted the body's response to each factor [3].

Table 1. – Influence of certain factors on physiological parameters [3]

Physiological	**Parameter**	**Mental Workload**	**Stress**	**Attention**	**Situational awareness**
ECG	Heart Rate	Increase	Increase	Increase	Increase
	HRV	Decrease	Decrease	Decrease	Decrease
RSP	Respiratory Rate	Increase	Increase	NA	NA
EDA	EDA	Increase	Increase	NA	NA
Eye tracking	Pupil diameter	Increase	NA	NA	NA
	Blink frequency	Decrease	Frequency increase	Decrease	Decrease
Task Related	Predefined tests/ Checklists/ Questionnaire/ Observations/ Recordings/ etc.	NASA-TLX test, reaction time, number of errors, accuracy	NA	Reaction time, Target detection, SART, Number of errors, Direct movement tracking	SART, SARS, SAGAT, SPAM, SALIANT, GIM

2. RELATED WORK

Ever since more intelligent and durable systems began to be used in many industries, we have seen safer and smarter equipment, from the DaVinci robot [4] for remote and precision surgery to car driving and self-braking assistance in the automotive field as well as much more advanced autopilot and safety features on airplanes. The only thing that did not change much is the human. Human factors have been studied across many domains ranging from human machine interaction to medicine. They will always be present in this kind of activities where decisions have to be made, thus it is necessary to study whether the person is able to manage situations and or the situation overwhelms the person.

The human performance envelope is not a new concept [11]-[18]. Physiology measurements on pilots were made even before 1975, one of the examples being "Heart rate monitoring of pilots during steep-gradient approaches" by Roscoe A. H. [5] and after that, many more followed. One of the early mentions of the Human performance envelope was in "Human Performance in air traffic control", Edward Tamsyn's PhD thesis [6], further cited in 2018 by Future Sky Safety. Although individual performance markers and factors were studied in the past, Edward's thesis is the first to bring together 9 factors that contribute to performance and illustrate their interdependencies. The 9 factors are: stress, mental workload, attention, situational awareness, fatigue, vigilance, communication, teamwork, and trust. These concepts were originally purposed/ intended for ATM. The envelope is represented by a spider graph in which the factors represent the 9 corners and a graphical illustration of performance is obtained. The illustration has a format that would very much favor the identification of the pilot's status whether his performance is optimal, acceptable, or degraded [7].

3. MATERIALS AND METHODS

The measurements were performed on the Boeing 737-Next Generation fixed-based flight simulator. The technology is a full-scale replica of the B737-NG aircraft cabin that artificially replicates aircraft flight and various aspects of the environment. The simulator is composed of specific software and hardware that provide a realistic view out of the cockpit and meets the requirements of the FNPT II simulator following EASA CS-FSTD (A) standards.

Fig. 2 - Flight simulator B737-NG during testing

Participants were asked to fly five scenarios using the simulator. Each scenario involved takeoff, climb, cruise, descent, approach, and touchdown phases of a flight and lasted approximately one hour.

A. Preflight preparation on the flight simulator

The chosen route is LROP Runway 08R to LRTR Runway 29, illustrated below in Fig. 3. Flight preparations included FMC programming.
The flight includes five waypoints with designated speeds and altitudes:

- **Departure SID:** SOKRU1A
- **Then waypoints:** ENIMA -NERDI- AGNEP
- **Arrival STAR:** NEKUL2A – note arrival was shortened from NEKUL direct to tactical fix TR221 assuming ATC clearence.

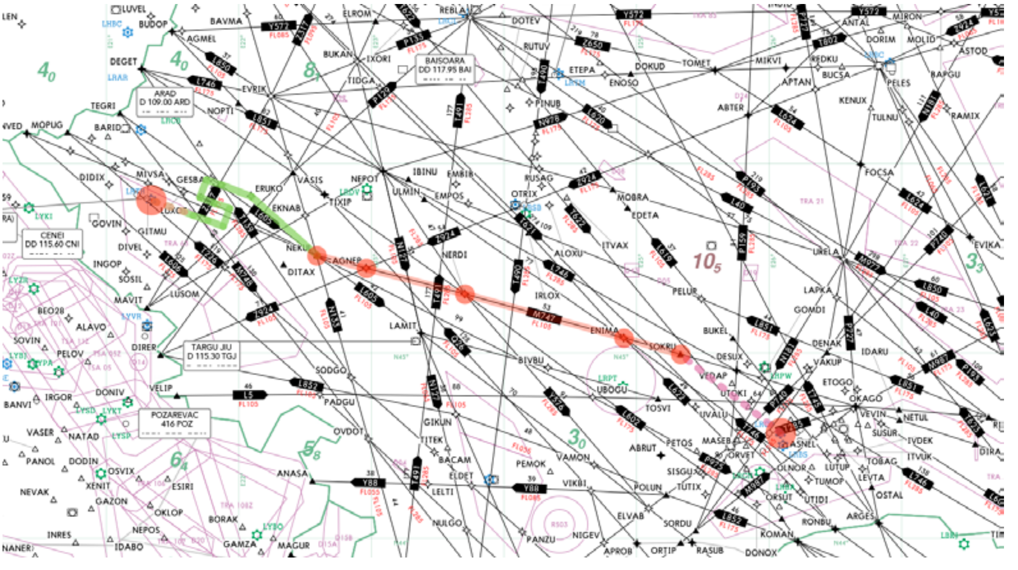


Fig. 3 - Full Flight Plan

B. Biopac System

The experimental environment included a Biopac MP160 using smart amplifiers to record physiological parameters. We paired the Biopac with a Tobii Pro Glasses 2 to also measure the pilot’s eye tracking data.

To record and preprocess the data, we used Biopac’s own software, AcqKnowledge and for the eye tracking metrics we used Tobii Pro Lab.

The smart amplifiers we used are specialized for electrocardiogram, electrodermal, respiration and electroencephalography.

While each smart amplifier has its own purpose, its filters can be turned off so that it can be repurposed for any other use with the adequate filters. Inherently, they have the basic filters already implemented like the 50Hz notch filter to rid the signal of the power grid noise (50 Hz in Europe & 60Hz North America) [8].

Starting from the D-II lead of the ECG signal registered from the pilot, we used a moving window, and we computed the positive envelope. A smoothing function was used on the obtained envelope and afterwards we identified the peaks representing breathing. The peaks were counted considering the sampling frequency and a respiratory rate per minute was obtained.

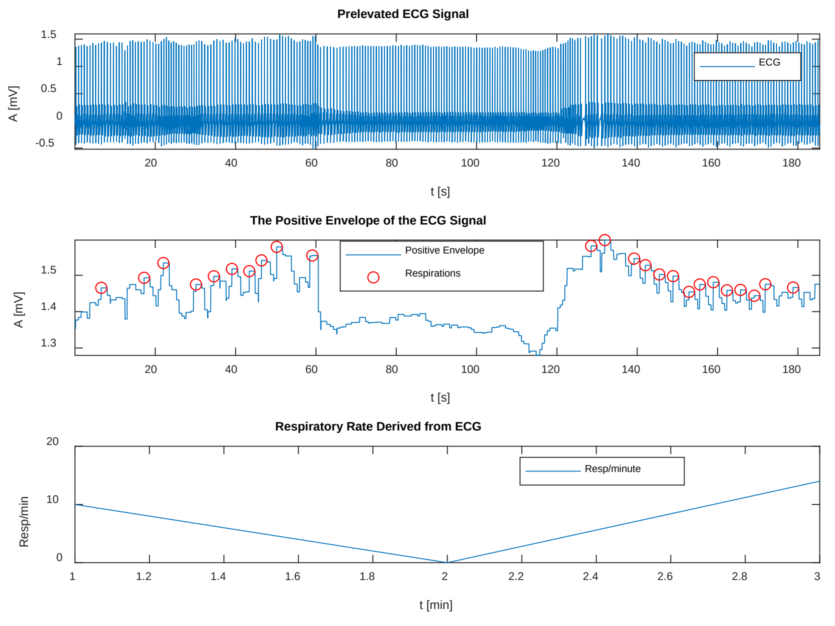


Fig. 4 - Respiratory rate computed from ECG

During the recording, the participant breathed normally while sitting for a minute, held his breath for a minute and then normal breathing for another minute. The above figure clearly illustrates the difference in the positive envelope while the participant held his breath, therefore there were no respirations detected during that minute.

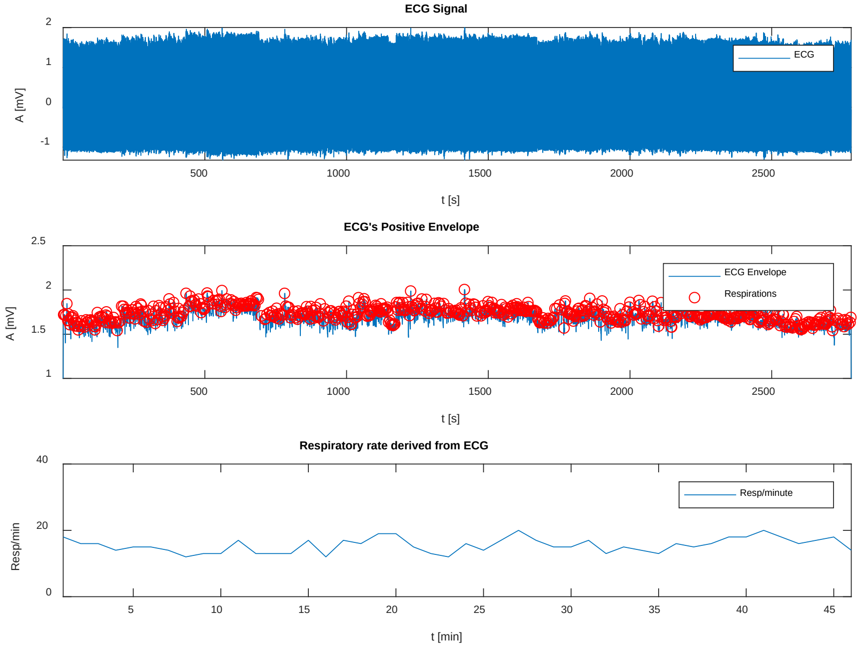


Fig. 5 - The algorithm was used for the ECG data recording of a full flight

Heart rate variability represents the variation in time between heartbeats, more specifically it is determined by measuring the variation in the RR intervals. It is influenced by the sympathetic and parasympathetic nervous system, therefore within the factors affecting it are sleep-wake cycle, physical activity and stress [7]. In our case we used the square root of the mean squared difference of successive RRs (RMSSD):

$$RMSSD = \sqrt{\frac{1}{N-1}\left(\sum_{i=1}^{N-1}(R-R)_{i+1}-(R-R)_i\right)^2} \tag{1}$$

where N represents the number of RR interval terms, R is the wave, and R-R is the interval between two consecutive R waves [8].

Considering that HRV is computed as multi-epoch HRV statistical with the epoch width of 60 seconds, it is best to compute the average values in every minute of flight for every physiological parameter. The increase or decrease of the data could provide information about changes in HPE's factors.

An increase of respiratory rate and HR along with the decrease in HRV show an increased level of workload or stress.

A sudden increase of HRV with a decrease of heart rate and respiratory rate would basically indicate a relaxation. In the figure below just before the 45 minute mark, the lowest level of HRV is achieved in the moment of approach and touchdown, considering that an engine failure was introduced in that specific flight.

After that, a spike forms on HRV with a decrease of HR and respiration rate, signifying a moment of relaxation, a moment of relief.

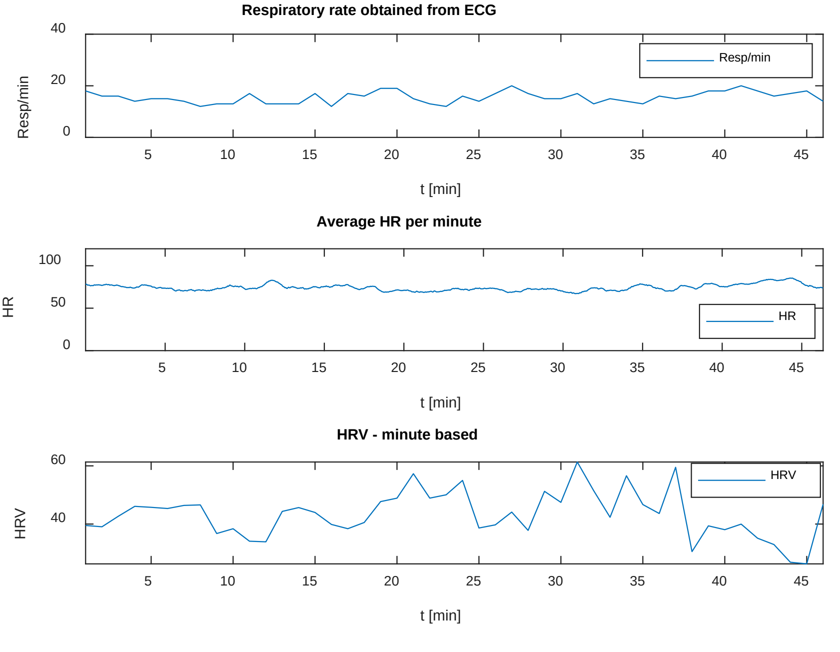


Fig. 6 - The figure illustrates the data we obtained and computed from the Biopac unit for a full flight. Around minute 37 of the recording an engine failure was introduced. While the first two signals don't show a clear change without in depth analysis, HRV dramatically changes

C. Eye Tracker

Eye tracking data were collected with the Tobii Pro Glasses 2. The system consists in 4 eye cameras, a high definition scene camera, a microphone, IR illuminators, a gyroscope and an accelerometer. The Glasses are connected to the Recording Unit via an HDMI cable. It has a simple one point calibration procedure and a 50Hz sampling frequency; the tracking technique used is Corneal reflection, binocular, dark pupil tracking. The pupil measurement is an absolute measurement [9].

In order to synchronize the data obtained from both systems, the Tobii Pro Glasses 2 were interfaced with the Biopac unit using a Biopac INISOA AMI/HLT. The connection is represented by an analogic 3.5mm wire that would send 3 consecutive pulses at the start of the recording of the glasses and then periodic pulses that would confirm the fact that the connection is still active.

Most of the recordings have had a data accuracy of over 93%. Data inaccuracy or errors are determined by the pilot looking outside the detection range of the glasses. Blinks were also generators of invalid data.

We tried to use these data to compute an estimated blink rate, excluding physically impossible samples. From its proprietary software we export the pupil data to Matlab for processing.

Pupil data processing consists in removing invalidities, computing an average pupil size for each minute of recording in order to match the Biopac data and obtaining an estimated blink rate that would also be calculated and represented per minute.

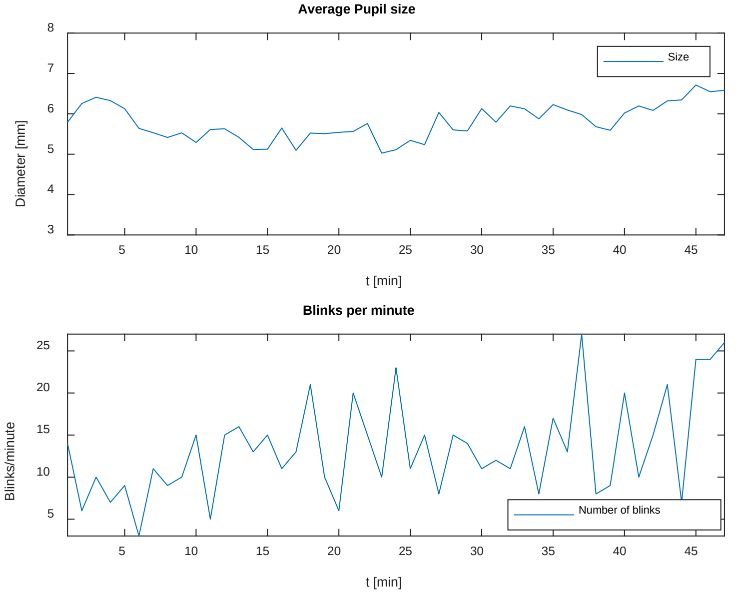


Fig. 7 - Average pupil size as exported from Tobii and calculated in Matlab. Blink rate determined in Matlab

An increase of pupil size or a decreased blink frequency can indicate a change in the factors that affect the subject. More exactly, these would be markers for an increase in workload. On the other hand, an increase of blink frequency could indicate the fact that stress levels have increased. Using Tobii Pro Lab's assisted mapping procedure we obtained a heatmap. The B737-NG cockpit has been divided into eight Areas of Interest (AOI) defined in Tobii Studio for analyzing the visual attention on the flight instruments.

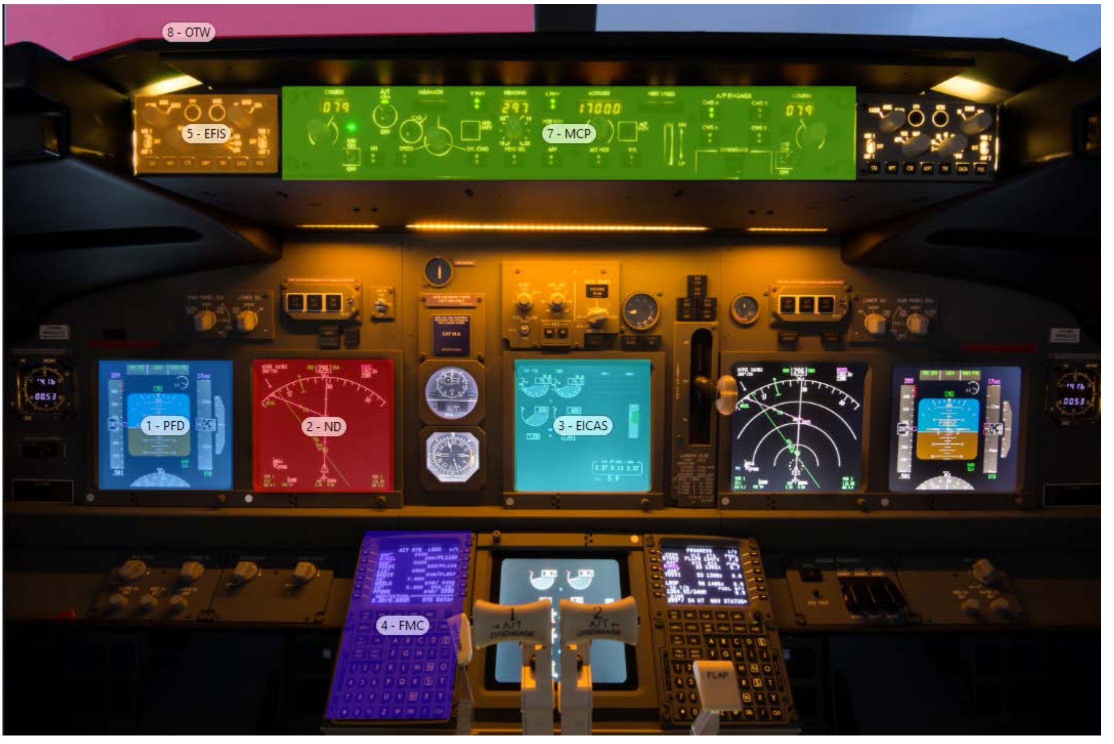


Fig. 8 - Cockpit Display with Areas of Interest: 1 - Primary Flight Display (PFD), 2 - Navigation Display (ND), 3 - Engine Indications and Crew Alerting System (EICAS), 4 - Flight Management Computer (FMC), 5 - Electronic Flight Instrument System (EFIS), 6 - Mode Control Panel (MCP), 7 - Out of Window (OTW)

D. NASA Task Load Index

NASA Task Load Index (NASA TLX) is a workload rating/assessment tool that consists of a questionnaire taken at different moments of a task from a wide range of applications.

Based on an average of ratings on Mental Demand, Physical Demand, Temporal Demand, Performance, Effort and Frustration, NASA TLX provides an overall workload score [10]. As a first notice, the mental demand was the highest, followed by the temporal demand.

This means that the tasks required a high level of concentration and had to be done in a short time.

E. Computing performance

Starting from the recorded physiological parameters and the subjective method presented above, a HPE model was shaped.

The following HPE model has three factors, all three from the group of factors directly linked to physiological readings.

A more complete HPE would contain up to nine factors that did not represent the scope of this study, either requiring extensive amounts of time for fatigue or psychological analysis like trust, communication, and teamwork.

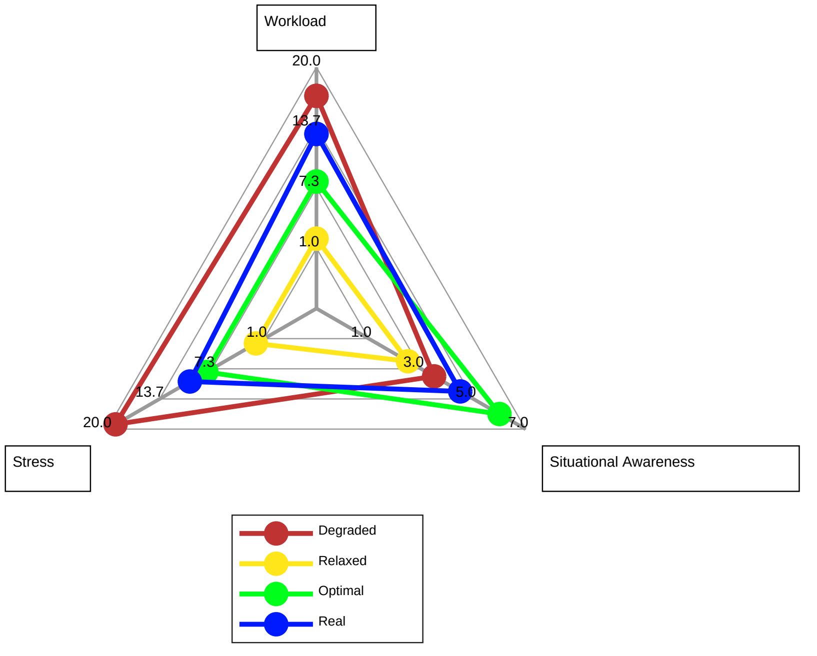


Fig. 9 - Computed model for HPE factor envelope

Considering they are interdependent [6], stress and workload influence situational awareness. It can be noticed that in the case of a very high or very low workload, the situational awareness is lowered.

For example, the relaxed state is representative for cruising where the pilot is just "supervising" the autopilot, both his workload and stress are low but at the same time the situational awareness is decreased due to the relaxed state.

In an optimal case, the pilot has a moderate amount of workload without much stress and has a high situational awareness.

The reality proves that the parameters are close to optimal in normal situations but, as soon as errors are introduced, they may degrade to dangerous levels.

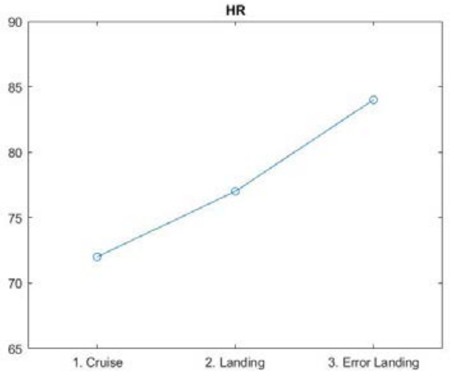
4. RESULTS

As shown in fig. 10 a – 10 e, there are significant changes in pilot's physiological parameters in more intense situations.

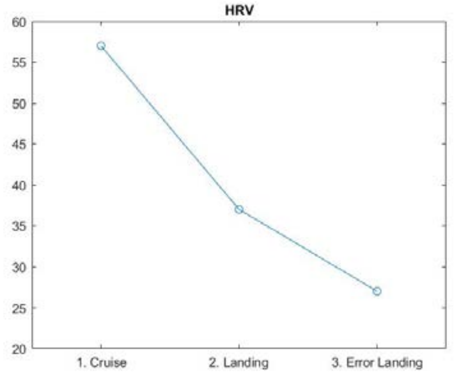
Cruising is considered a low effort task while the landing is the most difficult task in a flight. In the case of an error introduced for landing, the challenge was even higher.

The difficulty increase can be noticed in the results obtained by algorithms. The pilot's heart rate increased significantly, the heart rate variation decreased and/while the respiration rate increased.

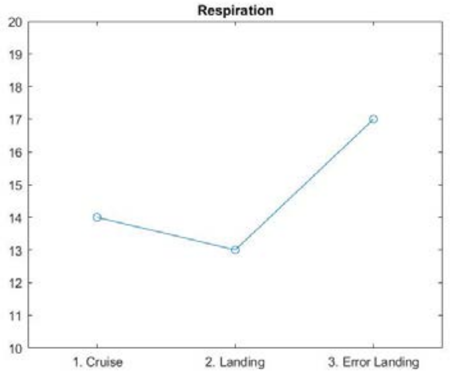
While pupil size increased, blink rate increased for a more difficult task but decreased as the pilot gazed out the window in order to land.



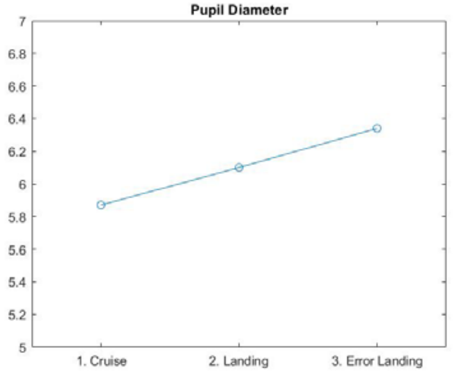
10 a



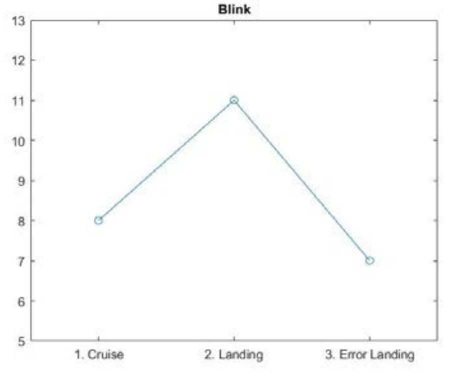
10 b



10 c



10 d



10 e

Fig. 10 a - Pilot's average heart rate and
b - The heart rate variation, in specific moments of the flight
c - Pilot's respiration rate
d - Pupil diameter and
e - Calculated blink rate

Below is a comparison of a heatmap from a normal landing (left) and a landing with an engine failure (right).

Relevant details about the pilot's attention can be determined from the heatmap. In the landing with the engine failure the pilot looked at the engine characteristics longer than in a normal flight, also, the pilot looked outside a lot more in order to successfully land.

Fig. 11 a - Heatmap of a normal landing

Fig. 11 b - Heatmap of a landing with an engine failure

Analyzing the flights, considering the average of the parameters per flight phase in each flight, in most of the flights, the results are appropriate, but it can be noticed that in the first flight there was a trim error that increased the workload and in the fifth flight due to an engine failure in landing, the workload was a little higher.

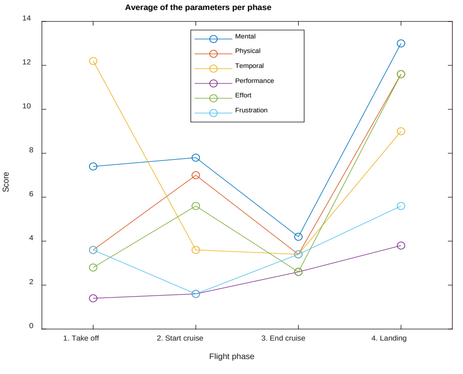


Fig. 12 - The average of the parameters per flight phase in each flight

Looking at the data, considering the average of the factors per flight phase, at takeoff, the temporal and mental demand were the highest. Also, in the next phase, the start cruise, the mental and physical demand were the highest. Towards end of the cruise, all factors were low and in the landing phase, the mental, effort and temporal demand were high. Also, the performance is inversely proportional, so even though it might look that the performance is low, actually the task has been accomplished adequately.

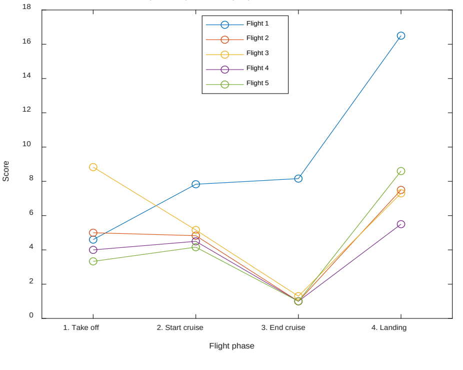


Fig. 13 - The average of the factors per flight phase

The Pupil Diameter has been examined using the Eye Tracker Tobii Pro Lab 2 and an Analysis of Variance (ANOVA) has been conducted. There were five flights analyzed and the overall average was 5.73 and the standard deviation was 0.15. Analyzing the data ($F_{4,277}$ = 13.14, $p < 0.05$) we can conclude that there is a significant level of variance between the pupil diameters in the sample. In other words, there was a variance in the sample, because of the errors introduced in some of the experimental sessions.

For the Heart Rate Variation, the data has been collected using Biopac and a rmANOVA has been conducted, using RMSSD and pNN50. Analyzing the RMSSD and pNN50 data ($F_{4,450}$ = 47.55, $p < 0.05$) we can conclude that there is a significant level of variance in the data.

During the five flights, the NASA TLX questionnaire has been concluded. The questionnaire has been taken in four parts of the flight. Using descriptive statistics, the average WL has been calculated. It can be observed that as the task difficulty increases, the WL increases.

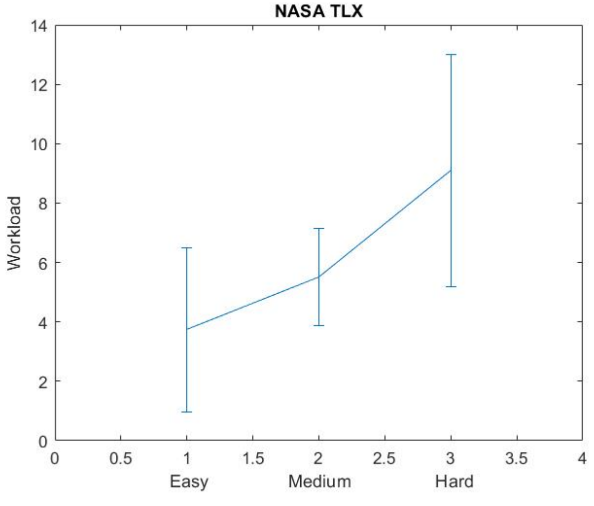


Fig. 14 - The Workload considering the difficulty of each task according to NASA TLX questionnaire

5. CONCLUSIONS

Our focus was on the factors that can be determined from physiological measurement such as Workload, Stress and Situational Awareness, without taking in consideration more psychological factors like communication, teamwork and trust.

It is noticeable that there are significant differences in both subjective WL and in the measured physiological parameters depending on the difficulty of the task. More difficult flight phases or a flight in which an error was introduced yielded higher levels of calculated and subjective WL. Stress is highly tied to workload, its physiological effects being similar. Situation Awareness is at its highest in medium levels of workload and stress. If the later are either too small or too big, SA can be greatly reduced, leading to long response times from the pilot in case of an incident. Unlike the before mentioned factors, attention requires other methods to be measured. An indicative for attention can be the heatmap, showing how the pilot's attention shifts depending on the situation.

The physiological parameters like the heart rate, respiratory rate and pupil diameter will increase with an increase of difficulty while the heart rate variation will decrease. The blink rate is more subjective to the task, an increase of difficulty would increase the blink rate, but the landing with an engine failure determined the pilot to gaze a lot longer on the outside window with a lower blink rate.

The Human Performance Envelope model gets more accurate as more physiological parameters are used. In our future studies we plan to include the electro-dermal response; for this it is necessary to find a spot on the subject's skin that would not interfere with piloting, also we have to take into consideration that a significant portion of the population does not have a clear response on the EDA. We will also add to our tests the Situation Awareness Rating Technique and other self-assessment questionnaires depending on the factors considered. To determine factors like fatigue and vigilance, longer tests are required in order to obtain a noticeable result.

REFERENCES

[1] * * * Bureau of Aircraft Accidents Archives, *www.baaa-acro.com*

[2] * * * Statistical Summary of Commercial Jet Airplane Accidents Worlwide Operations 1959-2016, Boeing.

[3] S. Silvagni, L. Napoletano, I. Graziani (DBL), P. Le Blaye (ONERA), L. Rognin (ECTL), *Concept for Human Performance Envelope*, Future Sky Safety, FSS_P6_DBL_D6.1-Concept-for-Human-Performance-Envelope_v2.0.pdf.

[4] * * * DaVinci Robot Surgery, *https://www.intuitive.com/en-us/products-and-services/da-vinci/systems*

[5] A. H. Roscoe, Heart rate monitoring of pilots during steep-gradient approaches, *Aviat Space Environ Med.*, 1975 Nov, **46**(11):1410-3.

[6] E. Tamsyn, *Human performance in air traffic control*, PhD thesis, University of Nottingham 2013, http://eprints.nottingham.ac.uk/id/eprint/28623.

[7] A. Martins et al., *Test report preliminary testing with system pilots cognitive task analysis*, Future Sky Safety, FSS_P6_DLR_D6.3_v2.0.pdf.

[8] * * * BIOPAC, *https://www.biopac.com/*

[9] * * * Tobii Pro Glasses 2, *https://www.tobiipro.com/product-listing/tobii-pro-glasses-2/*

[10] * * * NASA-TLX, *https://humansystems.arc.nasa.gov/groups/TLX/*

[11] T. Edwards, S. Sharples, J. R. Wilson, and B. Kirwan, *The need for a multi-factorial model of safe human performance in air traffic control*, no. August, pp. 25–27, 2010, Conference on Cognitive Ergonomics (pp. 253-260), ACM, DOI: 10.1145/1962300.1962353, (Edwardsetal.Multi-factorialmodelofsafehumanperformanceECCE.docx)

[12] H. Kim, E. Cheon, D. Bai, Y. H. Lee, and B. Koo, Stress and Heart Rate Variability: A Meta-Analysis and Review of the Literature, *Psychiatry Investig*, 2018 Mar, **15**(3):235-245, doi: 10.30773/pi.2017.08.17, Epub 2018 Feb 28., 2018.

[13] S. Nählinder, *Flight Simulator Training: Assessing the Potential*, Linköping Studies in Science and Technology Dissertations, No. 1250, 2009.

[14] I. Graziani, B. Berberian, B. Kirwan, P. Le Blaye, L. Napoletano, L. Rognin, S. Silvagni, *Development of the Human Performance Envelope Concept for Cockpit HMI Design*, HAL Id: hal-01409075, https://hal.archives-ouvertes.fr/hal-01409075, Submitted on 5 Dec 2016, DCSD16004.1452861830-1.pdf

[15] P. H. Charlton, T. Bonnici, L. Tarassenko, D. A. Clifton, R. Beale, P. J. Watkinson, An assessment of algorithms to estimate respiratory rate from the electrocardiogram and photoplethysmogram, *Physiol Meas.* 2016: **37**(4):610-626, doi:10.1088/0967-3334/37/4/610.

[16] F. H. Previc et al., The Effects of Sleep Deprivation on Flight Performance, *Instrument Scanning, and Physiological Arousal in Pilots*, vol. **19**, no. 4, pp. 326–346, 2009.

[17] K. K. E. Ellis, *Eye tracking metrics for workload estimation in flight deck operations*, University of IowaIowa Research OnlineTheses and Dissertations, University of Iowa, 2009.

[18] W.-C. Li and all, *The Investigation of Visual Attention and Workload by Experts and Novices, in the Cockpit the Investigation of Visual Attention and Workload by Experts and Novices in the Cockpit*, Proceedings of the 10th international conference on Engineering Psychology and Cognitive Ergonomics: applications and services - Volume **Part II**, DOI: 10.1007/978-3-642-39354-9_19, 2013.