Effectiveness of Flight Time Limitation (FTL)

D1 *Definition of the Baseline*

MOVE/C2/2016-360
Classification: Public

Table of Contents

Chapter 1: Introduction to the research study

Background

In commercial aviation, crew schedules are regulated by duty time limits, flight time limits, minimum rest requirements and other constraints. These limits and requirements, collectively referred to as flight time limitations (FTLs), are intended to be a simple method of limiting and accounting for fatigue among flight and cabin crew members, as part of overall safety concerns and objectives.

Over time, FTLs have evolved, driven by industrial pressures, new scientific data and the need to adapt to evolving aircraft capabilities. Nowadays, there are major differences among FTLs formulations in different parts of the world, influencing crew productivity and crew alertness, and operational flexibility. In view of the over-arching importance of the issue, there has been considerable research effort devoted in recent years on increasing the scientific knowledge and information in the areas of fatigue and alertness. The availability of such new research on sleep and work-related fatigue makes it ever more relevant to compare prevailing regulations with the new insights.

The Regulation (EC) No. 1899/2006[1] on the harmonisation of the technical requirements and administrative procedures in the field of civil aviation required the European Aviation Safety Agency (EASA) to conduct a scientific and medical review of Subpart Q of Annex III of the Regulation. The review was performed in 2008 and published in a Final Report titled "Scientific and Medical Evaluation of Flight Time Limitations". The evaluation of scientific studies in the field of fatigue management existing at that time revealed the need to conduct additional research to assess the quality of the new EU crew member fatigue management framework[2]. Comitology and Parliamentary scrutiny resulted in an instruction to EASA to perform a continuous review of the effectiveness of the provisions concerning flight and duty time limitations and rest requirements contained in Annexes II and III of Commission Regulation (EU) No. 965/2012[3]. This mandate was formalised in Article 9b of this Regulation.

Main objective and scope of the research study

The review commenced in 2017 with the commission of a research study. The review includes notably an assessment of the impact of at least the following on the alertness of aircrew:

1. Duties of more than 13 hours at the most favourable time of the day;
2. Duties of more than 10 hours at the less favourable time of the day;
3. Duties of more than 11 hours for crew members in an unknown state of acclimatisation;
4. Duties including a high level of sectors (more than six);
5. On-call duties such as standby or reserve, followed by flight duties; and
6. Disruptive schedules.

[1] Regulation (EC) No. 1899/2006 of the European Parliament and of the Council of 12 December 2006 amending Council Regulation (EEC) No. 3922/91 on the harmonisation of technical requirements and administrative procedures in the field of civil aviation.

[2] Commission Regulation (EU) No. 83/2014 of 29 January 2014 amending Regulation (EU) No. 965/2012 laying down technical requirements and administrative procedures related to air operations pursuant to Regulation (EC) No. 216/2008 of the European Parliament and of the Council.

[3] Commission Regulation (EU) No. 965/2012 of 5 October 2012 laying down technical requirements and administrative procedures related to air operations pursuant to Regulation (EC) No. 216/2008 of the European Parliament and of the Council.

The review was broken down into smaller research phases; each focused on specific flight duty periods (FDPs). The first and current research phase studied the two types of duty expected to pose the highest level of fatigue.

Phase 1: Specific objectives and technical tasks

The current FTL study included the following work content:
- Ranking the chosen aircrew duty periods based on the expected level of fatigue and selection of the two that are top ranking;
- Identification of a representative population and a relevant type of operations to be used for data sampling purposes;
- Detection of potential fatigue hotspots in commercial air transport (CAT);
- Collection of objective and subjective data on mental effort, fatigue and performance for a target aircrew population;
- Benchmarking of the study with other relevant sources;
- Assessment of the fitness-for-purpose of the regulatory fatigue management controls to ascertain the need for any additional mitigations deemed necessary; and
- Drawing conclusions on the work performed and drafting recommendations.

Scope of the current deliverable

This Deliverable D1 (Definition of the Baseline) reports the results of the work performed on the first two tasks/bullets above; i.e., the ranking of the six duty periods of interest and the identification of a representative set of air transport operators. D1 starts with a review of the state of the art.

Chapter 2: Review of the state of the art

A review of the state of the art was performed by collecting the required background information that is relevant for establishing a robust and detailed baseline – viz. the starting point – for the research study.

Critical assessment of existing references

The review of the state of the art included a critical assessment of existing research relevant to aviation fatigue and alertness with a view to collect background information on data collection activities.

In view of the broad scope, complexity and sensitivity of the intended data collection, particular attention was given to highlight 'lessons learned' from previous campaigns in view of smoothing any difficulties that might arise in the future data gathering exercise.

Literature review

A search was conducted in the different search engines such as PubMed and Google Scholar, using different combinations of – or variations of – the search terms fatigue, sleep(iness), drowsiness, alertness, flight/duty time limitation, shift/night work, time zone changes, jet lag, air transport/operations, flight-/aircrew, pilot, and cabin crew. Additional references were identified by manually searching bibliographies of the retrieved publications. Furthermore, scientists from within the project team's network were requested to provide recent publications on aircrew measurement campaigns on fatigue and alertness.

The publications examined were systematically assessed with emphasis on the following topics:

- Selection population (*who was being measured?*);
- Measurement techniques (*what was being measured?*);
- Scale of data collection (*how many participants were being measured and for how long, in what geographical region?*);
- Protocols followed (*when and how were measurements taken?*); and
- Objectives and conclusions (*what was being studied and concluded?*).

The suitability of the selected data sources to be used for purposes of benchmarking of the results was also assessed. The publications were categorised into two groups, the first being the most relevant for the benchmark as these publications 'fit best' to the current research:

1. Fatigue-related research studies AND European commercial aviation (i.e., operations undertaken under EU regulations) AND published after 2006; and
2. Fatigue-related research studies AND non-European commercial aviation AND published after 2006.

A cut-off regarding publication date was used to contain the extent of this state of the art exercise. Furthermore, publications before 2007 were covered by the review "Scientific and Medical Evaluation of Flight Time Limitations".

Simulator studies were not included in the review of the state of the art as the focus was on aircrew measurement campaigns in an operational setting, similar to the future data gathering exercise.

Lessons learned

Valuable lessons can be learned from previous crew fatigue studies on how to smoothly collect a series of measurements to provide high quality data. Therefore, in addition to the literature search, lessons learned in previous studies were systematically identified and documented. The lessons considered technical (e.g. which type of measures to use?), organisational (e.g. how to ascertain that the equipment is available and the data collected is retrieved correctly?), and personal aspects (e.g. how to make sure that subjects remain motivated to participate in the study?).

The lessons learned were gathered via semi-structured interviews with senior researchers and project managers that were actively involved in crew fatigue studies. Airlines that undertook or participated in fatigue studies were also contacted to acquire lessons learned from their perspective.

Mapping of collected state of the art

The literature search and the gathering of the lessons learned resulted in a mapping of the collected state-of the art that is perceived as relevant.

Literature overview

The literature search resulted in 34 relevant publications that were included. These publications were classified into one of the two groups (Group I: EU and Group II: non-EU), resulting in the following mapping of the publications (Table 1).

Table 1 Mapping of collected publications

Author(s)	***Year***	***Group***	***Who is being measured***	***Which variables are being measured***	***Number of participants***	***Geographical region***	***Type of operation***
Cabon et al.	2012	I. EU	Pilot(s)	FDM, ASR	N = 230	Europe	Regional
Houston et al.	2012	I. EU	Aircrew	Fatigue Report	N = 983	Europe	
Reis et al.	2013	I. EU	Pilot(s)	FSS	N = 456	Europe	Mixed
Brown et al.	2014	I. EU	Aircrew	KSS, SP, PVT, actigraphy	N = 14	Europe	Mixed
Ingre et al.	2014	I. EU	Pilot(s)	KSS, sleep log	N = 136	Europe	Mixed
Vejvoda et al.	2014	I. EU	Pilot(s)	Actigraphy, KSS, sleep log, SP, NASA TLX	N = 40	Europe	Short-haul
Reis et al.	2016	I. EU	Pilot(s)	FSS, JSS, ESS	N = 435	Europe	Mixed
O'Hagan et al.	2016	I. EU	Pilot(s)	Questionnaire	N = 954	Europe	
Sallinen et al.	2017	I. EU	Pilot(s)	KSS, actigraphy, sleep log	N = 90	Europe	Mixed
Srivistava & Barton	2012	I. EU	Aircrew	PVT, SP, sleep log, ESS, actigraphy, NASA TLX	N = 22	Europe	Short-haul
Powell et al.	2007	II. Non-EU	Pilot(s)	SP	2034 duties	New Zealand	Short-haul
Thomas et al.	2007	II. Non-EU	Pilot(s)	PVT, actigraphy, sleep log	N = 37	Australia	
Powell et al.	2008	II. Non-EU	Pilot(s)	SP	4206 duties	New Zealand	Short-haul
Mello de et al.	2008	II. Non-EU	Pilot(s)	FDM, ASR	N = 987	South-America	

Powell et al.	2010	II. Non-EU	Pilot(s)	KSS, SP, PVT	N = 24/21/27	New Zealand	Long-haul
Roma et al.	2010	II. Non-EU	Cabin crew	PVT, actigraphy, sleep log, KSS	N = 202	North America	Regional
Thomas & Ferguson	2010	II. Non-EU	Pilot(s)	TEM	302 flight ops	Australia	Short-haul
Powell et al.	2011	II. Non-EU	Pilot(s)	SP	4629 ratings	New Zealand	Short-haul
Roach et al.	2011	II. Non-EU	Pilot(s)	Sleep log, actigraphy	N = 301	Australia	Long-haul
Roach, Sargent et al.	2012	II. Non-EU	Pilot(s)	SP	N = 70	Australia	Short-haul
Holmes et al.	2012	II. Non-EU	Pilot(s)	KSS, actigraphy, sleep log	N = 44	Middle East	Long-haul
Roma et al.	2012	II. Non-EU	Aircrew	PVT, actigraphy, activity/sleep log	N = 201	North America	
Roach, Petrilli et al.	2012	II. Non-EU	Pilot(s)	PVT, actigraphy, sleep log, SP	N = 19	Australia	Long-haul
Gander, Signal et al.	2013	II. Non-EU	Pilot(s)	Actigraphy, sleep log, KSS, SP, PVT	N = 70	New Zealand	Long-haul
Wu	2013	II. Non-EU	Pilot(s)	PVT, actigraphy, sleep log, SP	N = 74	North America	Long-haul
Greeley et al.	2013	II. Non-EU	Cabin crew	PVT, speech recording	N = 195	North America	Long-haul
Gander, van den Berg et al.	2013	II. Non-EU	Pilot(s)	Actigraphy, sleep log, KSS, SP, PVT	N = 30	New Zealand	Short-haul
Signal et al.	2013	II. Non-EU	Aircrew	Actigraphy, sleep log, polysomnography	N = 21	New Zealand	Long-haul
Gander, Mangie et al.	2014	II. Non-EU	Pilot(s)	Actigraphy, sleep log, KSS, SP, PVT	133 landings	New Zealand	Long-haul
Gander, Mulrine et al.	2014	II. Non-EU	Pilot(s)	Actigraphy, sleep log, KSS, SP, PVT	N = 237	North America	Long-haul
Gander et al.	2015	II. Non-EU	Pilot(s)	Actigraphy, sleep log, KSS, SP, PVT	N = 237	North America	Long-haul
Berg van den et al.	2016	II. Non-EU	Pilot(s)	KSS, SP	N = 586	New Zealand	Long-haul
Signal et al.	2014	II. Non-EU	Pilot(s)	Actigraphy, sleep log, KSS, SP, PVT	N = 52	New Zealand	Long-haul
Berg van den et al.	2015	II. Non-EU	Cabin crew	Actigraphy, sleep log, KSS, SP, PVT	N = 55	New Zealand	Long-haul

ASR: air safety report; ESS: Epworth Sleepiness Scale; FDM: flight data monitoring; FSS: fatigue severity scale; JSS: Jenkins Sleep Scale; KSS: Karolinska Sleepiness Scale; NASA TLX: NASA task load index; PVT: psychomotor vigilance task; SP: Samn-Perelli; TEM: threat and error management.

Group I: Fatigue-related research studies in EU (after 2006)

Ten publications were included in Group I. The same research study[4,5] was published twice in different journals; this accounted for one research study. The NASA – easyJet collaboration on the human factors monitoring program[6], resulting in several interim reports, was also seen as one research study.

[4] Reis, C., Mestre, C., Canhão, H., Gradwell, D., & Paiva, T. (2016). Sleep and fatigue differences in the two most common types of commercial flight operations. *Aerospace Medicine and Human Performance*, *87(9)*, 811-815.

[5] Reis, C., Mestre, C., Canhão, H., Gradwell, D., & Paiva, T. (2016). Sleep complaints and fatigue of airline pilots. *Sleep Science*, *9(2)*, 73-77.

[6] Srivistava, A. S., & Barton, P. (2012). Collaboration on the human factors monitoring program (HFMP) study. NASA Report No. TM-2012-216053.

Straight-forward observations of the included publications showed the following:

- All ten studies included pilots and two studies also included cabin crew. The type of operation studied was not clearly identified in two studies; five studies concerned a mix of short- and long-haul operations; two studies were on short-haul operations; and one on regional operations;
- Regarding the measurements, three out of ten used paper forms or questionnaires. The other seven measured digitally through web-based surveys, the use of personal digital assistants (PDAs) or smartphones. Only the studies using PDAs performed in-flight assessments; no paper-based studies were found assessing aircrew in-flight;
- A large set of different measures was used within the ten studies, ranging from flight data monitoring and air safety reports to several fatigue/sleep rating scales (FSS, SP, KSS, ESS, and JSS). Sleep logs and/or actigraphy was used in five studies;
- The number of participating aircrew ranged from n=14 to n=983. Studies, where participants were asked to rate fatigue/sleep or perform a vigilance test in-flight, did not exceed n = 90; and
- Four of the included publications studied the prevalence of fatigue or sleep complaints. One researched blue light as a fatigue countermeasure. Two studies were performed in support of the implementation of a fatigue risk management system (FRMS).

Out of the ten included studies, three showed high resemblance with the current FTL research study; i.e., regarding the research goals, likely target population (note that both pilots and cabin crew were only targeted in the Srivista & Bartson study), and likely measurements used. These studies were used as a primary source in the benchmark activities to come. The benchmark studies are:

1. Sallinen, M., Sihvolaa, M., Puttonena, S., Ketolac, K., Tuoric, A., Härmäa, M., Kecklundd, G., & Åkerstedt, T. (2017). Sleep, alertness and alertness management among commercial airline pilots on short-haul and long-haul flights. *Accident Analysis & Prevention*, *98*, 320-329.
2. Srivistava, A. S., & Barton, P. (2012). Collaboration on the human factors monitoring program (HFMP) study. NASA Report No. TM-2012-216053.
3. Vejvoda, M., Elmenhorst, E. M., Pennig, S. B., Parh, G., Maass, H., Tritschler, K., Basner, M., & Aeschbach, D. (2014). Significance of time awake for predicting pilots' fatigue on short-haul flights: implications for flight duty time regulations. *Journal of Sleep Research*, *23(5)*, 564-567.

Note that both Sallinen et al. and Vejvoda et al. referred to the EU FTL regulations No. 83/2014 in their discussions and related their results to these new regulations that are under review in the current research.

<u>Group II: Fatigue-related research studies in non-EU (after 2006)</u>

Twenty-four publications were included in Group II. An overview of what was learned from these articles is listed below:

- 21 studies included pilots, 2 of these studies also included cabin crew and there were three studies that only included cabin crew;
- In most studies multiple measures were collected. Most studies used rating scales as measurement instruments. Self-rated sleepiness was measured using the KSS in 12 studies, and self-rated fatigue was measured using the SP in 11 studies. To actually measure sleep quantity, actigraphy was used in combination with sleep logs in 15 studies. As indicators of performance the PVT was used in 14 studies. Some instruments were rarely used; i.e., TEM system and flight operations quality assurance (FOQA) program as a performance indicators, and speech recordings

were only used once in the selected studies. Some studies had as goal to validate measurement instruments and therefore used polysomnography to compare between instruments; and

- Seventeen studies were undertaken on operations based in Australia/New-Zealand, five based in North America, one based in South America, and one based in the Middle East. Eleven studies were relatively small, comprising 19 to 74 subjects. The biggest study comprised 987 subjects. Others expressed their size in number of filled in rating scales (the largest had 4629 responses) or flight operations (in total 302 operations). Thirteen studies focused specifically on long-haul, six on short-haul and one on both (mixed). One study focussed on regional flights within North America. The remaining studies did not indicate the type of operations.

From all 24 studies, the 15 studies listed below were of extra interest because the design and applied methods showed great resemblance with the approach that is foreseen for the current study. These studies were used as a secondary source (given the fact that these studies originate from outside Europe) in the benchmark activities.

1. Gander, P. H., Mulrine, H. M., van den Berg, M. J., Smith, A. A. T., Signal, T. L., Wu, L. J., & Belenky, G. (2014). Pilot fatigue: relationships with departure and arrival times, flight duration, and direction. *Aviation, Space, and Environmental Medicine*, *85(8)*, 833-840.
2. Gander, P. H., Signal, T. L., Berg, M. J., Mulrine, H. M., Jay, S. M., & Mangie, J. (2013). In-flight sleep, pilot fatigue and psychomotor vigilance task performance on ultra-long range versus long range flights. *Journal of Sleep Research*, *22(6)*, 697-706.
3. Greeley, H. P., Roma, P. G., Mallis, M. M., Hursh, S. R., Mead, A. M., & Nesthus, T. E. (2013). Field study evaluation of cepstrum coefficient speech analysis for fatigue in aviation cabin crew. FAA Report No. DOT/FAA/AM-13/19.
4. Holmes, A., Al-Bayat, S., Hilditch, C., & Bourgeois-Bougrine, S. (2012). Sleep and sleepiness during an ultra long-range flight operation between the Middle East and United States. *Accident Analysis & Prevention*, *45*, 27-31.
5. Mello, M. T. de, Esteves, A. M., Pires, M. L. N., Santos, D. C., Bittencourt, L. R. A., Silva, R. S., & Tufik, S. (2008). Relationship between Brazilian airline pilot errors and time of day. *Brazilian Journal of Medical and Biological Research*, *41(12)*, 1129-1131.
6. Powell, D. M. C., Spencer, M. B., & Petrie, K. J. (2010). Fatigue in airline pilots after an additional day's layover period. *Aviation, Space, and Environmental Medicine*, *81*, 1013-1017.
7. Powell, D., Spencer, M. B., Holland, D., & Petrie, K. J. (2008). Fatigue in two-pilot operations: implications for flight and duty time limitations. *Aviation, Space, and Environmental Medicine*, *79(11)*, 1047-1050.
8. Powell, D., Spencer, M. B., Holland, D., Broadkent, E., & Petrie, K. J. (2007). Pilot fatigue in short-haul operations: effects of number of sectors, duty length, and time of day. *Aviation, Space, and Environmental Medicine*, *78*, 698-701.
9. Roach, G. D., Petrilli, R. M. A., Dawson, D., & Lamond, N. (2012). Impact of Layover length on sleep, subjective fatigue levels, and sustained attention of long-haul airline pilots. *Chronobiology International*, *29(5)*, 580-586.
10. Roach, G. D., Sargent, C., Darwent, D., & Dawson, D. (2012). Duty periods with early start times restrict the amount of sleep obtained by short-haul airline pilots. *Accident Analysis & Prevention*, *45*, 22-26.
11. Signal, T. L., Mulrine, H. M., van den Berg, M.J., Smith, A. A. T., Gander, P. H., & Serfontein, W. (2014). Mitigating and monitoring flight crew fatigue on a westward ultra-long-range flight. *Aviation, Space, and Environmental Medicine*, *85*, 1199-1208.

12. Thomas, M. J. W., & Ferguson, S. A. (2010). Prior sleep, prior wake, and crew performance during normal flight operations. *Aviation, Space, and Environmental Medicine*, *81(7)*, 665-670.
13. Thomas, M. J. W., Petrilli, R., M., & Roach, G. D. (2007). The impacts of Australian transcontinental 'back of clock' operations on sleep and performance in commercial aviation flight crew. ATSB Report No. B2005/0121.
14. Berg, M. J. van den, Signal T. L., Mulrine, H. M., Smith, A. A. T., Gander, P. H., & Serfontein, W. (2015). Monitoring and managing cabin crew sleep and fatigue during an ultra-long range trip. *Aerospace Medicine and Human Performance*, *86(8)*, 705-713.
15. Wu, L. J. (2013). Evidence based fatigue risk management during 24/7 operations: objective assessment of pilots' sleep, performance, and fatigue during ultra long range and long range flights. Doctoral dissertation, Washington State University.

Gathered lessons

Eight interviews were held to gather lessons learned on aircrew fatigue/alertness field studies. The interviewees concerned six fatigue experts that had recently worked on large-scale studies in operational aviation environment; i.e., representatives of all four project partners (NLR, Stockholm University, DLR, and Jeppesen) and two subject matter experts (with an academic background) from outside the consortium were interviewed. In addition, two airline representatives (from UK and Germany) shared the best tips and tricks regarding these measurement campaigns from their point of view.

The interviews were held by telephone and each took about one hour to complete. The notes of the eight interviews were checked for lessons learned. There could be a lesson that was put forward by more than one interviewee, or a lesson that was addressed by a single interviewee with extra emphasis. Contradictory lessons were disregarded. The lessons identified in the interviews were grouped into different subjects. The following lessons learned were identified:

How to approach the participants?

- Participants should be approached through the airline company (e.g. internal portal, mailings, or through internal contacts) and the unions;
- To get aircrews to participate it is important to guarantee their anonymity; and
- Support of high-level management within an airline is also important.

How to motivate the participants?

- Incentives (e.g. a raffle) are used quite often (successfully) to motivate aircrews to participate;
- Participating aircrews like getting some feedback on study results at end of studies; however, this requires great effort and coordination from the research team, especially in large-scale studies;
- It is important to train participants adequately in using the equipment and applying the measurement protocol, and build a friendly, professional relationship with them; and
- Pilots are often technically adept, so if you give them a gadget to collect data, then they are naturally interested in it.

What about coordination?

- Within the airline there is the need for a primary (internal) contact or coordination team (preferably from operations or health/safety people) for the aircrews to contact if needed; and within the research organisation there is the need for a central point of contact for the airline coordination team to contact for support;

- It must be possible for aircrews to contact the researcher directly, without any involvement of the airline company. Emphasise the independence of the researchers from the airline; and
- There is a need to have face-to-face contact (not only e-mail contact) with the participants somehow. This can be via a coordinator on site. Think of more than one coordinator. These coordinators can be used for training, data handling, motivation, etc. Also do not forget to provide incentives/rewards for the coordinators. And the researchers should visit the coordinators as much as possible.

How to train participants?

- There are several scenarios possible with regard to the training of participants:
  - If there is just the use of questionnaires and/or actigraphs, send it by e-mail;
  - If some more equipment is required, brief the crew before flight at crew centre; and
  - If measurements are numerous/complex, arrange a separate time for a crew briefing and allow adequate time.

How to get high-quality data?

- Electronic data gathering (through hand-held devices) as opposed to paper/pencil seems to be greatly appreciated; i.e., by the participating aircrews and also by the researchers handling the data; and
- If the participant indicates to they are overburdened by the measures, advise them to drop the PVT first as this appears to be the biggest burden.

Which measurement techniques to use in operational setting?

- Actigraphs work very well in the operational setting, so do rating scales. Performance testing requires some training and motivation; the biggest problem here is that there is no control over test environment;
- Subjective sleepiness measurements have been shown to be very much influenced by surroundings; i.e., it is not clear if people are involved in a talk, travel, etc. Advice is to always ask for sleepiness in last five minutes, not ‘now’;
- People are really bad at reflecting on their (subjective) sleep quality; an actigraph should be used for this; and
- With the rating of e.g. KSS is it advised to not show the previous score to prevent bias.

What about missing data points?

- If the analysis is based on (mixed model) regression, missing data points are not that important.

Data collection in air operations – concluding remarks

The findings of the literature review and the interviews resulted in a number of issues that needed to be addressed in the (set-up of the) field data collection study.

Inclusion and training of participants

- It was highlighted by interviewees that the key to successful recruitment is having all interested parties familiar with the reasons for, and in agreement with, the data collection. Recruitment involves collaboration between research teams, airlines, union representatives, and aircrew. Consulting with each party during protocol development has helped facilitate study recruitment;

- Participation in research studies is voluntary. It is feasible that participation in the data collection could be incentivized. Studies (such as Ingre et al.[7]) that have used incentives showed good results regarding the recruitment of participants and the commitment of these participants to complete the measures throughout the entire length of the research period;
- Depending on the types of measures employed, training can be completed either in person or remotely using internet and web-conferencing, or by providing a computer-based training online. The latter is relevant in case of a large sample size and a straightforward measurement protocol; and
- It is advised that a dedicated team is formed per participating airline that coordinates the data collection campaign inside the airline. This team should consist of motivated personnel (preferably two people from a health or safety department) who know their way around within the airline. The team is completed by one or two representatives of the project consortium, fully informed on the measurement protocol. Participating crews should clearly know where to go to in case of issues or lack of clarity. In case of a somewhat smaller sample size and a more complicated measurement protocol, the coordination team should provide the required training. Note that it may also be of value to reward or incentivize the airline personnel in the coordination team to keep them involved and motivated.

Measurement techniques

Field studies often use sleepiness or fatigue ratings. One frequently used measure is the KSS, which measures sleepiness on a 9-point scale from extremely alert to extremely sleepy, fighting sleep. It has been validated[8] and is used to measure subjective sleepiness in both laboratory and field studies[9,10]. Level 7 indicates the start of electroencephalographic and electrooculographic changes representing sleepiness, and level 8 - 9 is associated with high probability of line crossings on real roads and accidents in simulators[11,12].

Another relevant and frequently used rating scale is the SP. SP is a 7-point scale with possible scores ranging from 1 ("fully alert, wide awake") to 7 ("completely exhausted, unable to function effectively"). The SP crew status check was developed specifically for use with flight crew[13,14]. It has been used in studies focused on sleep loss, fatigue, and performance of flight crew[15,16,17], as well as in laboratory studies[18]. On the SP

[7] Ingre, M., van Leeuwen, W., Klemets, T. et al. (2014). Validating and extending the three process model of alertness in airline operations. *PLoS One*, 9.

[8] Kaida, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y. et al. (2006). Validation of the Karolinska sleepiness scale against performance and EEG variables. *Clinical Neurophysiology*, *117*(7), 1574-1581.

[9] Gillberg, M., Kecklund, G., & Åkerstedt, T. (1994). Relations between performance and subjective ratings of sleepiness during a night awake. *Sleep*, *17(3)*, 236-241.

[10] Härmä, M., Sallinen, M., Ranta, R., Mutanen, P., & Muller, K. (2002). The effect of an irregular shift system on sleepiness at work in train drivers and railway traffic controllers. *Journal of Sleep Research*, *11(2)*, 141-151.

[11] Åkerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience*, *52*, 29-37.

[12] Åkerstedt, T., Anund, A., Axelsson, J., & Kecklund, G. (2014). Subjective sleepiness is a sensitive indicator of insufficient sleep and impaired waking function. *Journal of Sleep Research*, *23*, 240-252.

[13] Samn, S. W., & Perelli, L. P. (1982). Estimating aircrew fatigue: a technique with implications to airlift operations. USAF School of Aerospace Medicine Technical Report No. SAM-TR-82-21.

[14] Samel, A., Wegmann, H. M., Vejoda, M., Drescher, E. E. J., Gundel, A., Manzey, D., & Wenzel, J. (1997). Two-crew operations: stress and fatigue during long-haul night flights. *Aviation, Space, and Environmental Medicine*, *68(8)*, 679-687.

[15] Pascoe, P. A., Johnson, M. K., Roberston, K. A., & Spencer, M. B. (1995). Sleep in rest facilities on board aircraft: field studies. DERA Report No. DERA/CHS/A&N/CR/95/002.

[16] Robertson, K. A., Spencer, M. B., Stone, B. M., & Johnson, M. K. (1997). Scheduling the on-board rest of aircrew. DERA Report No. DERA/CHS/PP5/CR97095/1.0.

there is less empirical evidence for a cut-off comparable to level 8-9 of the KSS. However, values of five and above have been used in previous studies as an indicator of fatigue in aviation operations[19,20,21].

Both the KSS and SP have the advantage of being easy to use in an operational environment. This is illustrated by the high number of times both measures are applied in the reviewed research studies.

Total sleep time is often measured in the field with actigraphy, a wrist-watch like device that uses an accelerometer to measure activity counts in a given time period. The activity count record is scored by a sleep scoring algorithm implemented in software to create a minute-by-minute sleep/wake history. Total sleep time measured by actigraphy is highly correlated with that measured by polysomnography among flight crews while in flight and during layover[22]. A number of actigraphs are available commercially. A device that measures activity in 60 second bins is appropriate for estimating total sleep time[23].

A strength of actigraphy is the ability to monitor individuals' activity levels over time, and thus to monitor sleep over a period of days to months. Actigraphs with adequate memory space can be worn for three months at a time before requiring download. Since the actigraph predicts sleep time based on activity, time periods where the user removes the device and places it in a stationary place may be scored as sleep. Newer devices offer off-wrist detection, using a sensor to determine if the device is being worn.

A sleep diary or sleep log is a record of an individual's sleeping and waking times, typically made over a period of several weeks. It is an inexpensive technique and provides a subjective estimate of sleep quantity and quality. Actigraphy and sleep diary data are often collected in parallel. Signal et al.[24] compared actigraphic and subjective estimates of flight crew sleep to the 'gold standard'. For estimating mean sleep duration, both actigraphic and subjective estimates were sufficiently close to polysomnographic values, but the amount of random error must be considered here. Any single estimate may vary by more than one hour from the mean difference. Van

---

[17] Samel, A., Wegmann, H.-M., & Vejvoda, M. (1997). Aircrew fatigue in long-haul operations. *Accident Analysis & Prevention*, *29(4)*, 439-452.
[18] Ferguson, S. A., Paech, G. M., Sargent, C., Darwent, D., Kennaway, D. J., & Roach, G. D. (2012). The influence of circadian time and sleep dose on subjective fatigue ratings. *Accident Analysis & Prevention*, *45*, 50-54.
[19] Powell, D. M., Spencer, M. B., & Petrie, K. J. (2011). Automated collection of fatigue ratings at the top of descent: a practical commercial airline tool. *Aviation, Space, and Environmental Medicine*, *82(11)*, 1037-1041.
[20] Gander, P., van den Berg, M., Jay, S., & Signal, T. (2012). Comparison of flight crew sleep and fatigue during Delta Air Lines long range and ultra-long range operations. Sleep/Wake Research Centre, Massey University.
[21] Berg, M. J. van den, Signal, T. L., Mulrine, H., Smith, A. A. T., & Gander, P. H. (2013). Evaluation of the sleep and performance of South African Airways cabin crew on the Johannesburg-New York ultra long range flight. Sleep/Wake Research Centre, Massey University.
[22] Signal, T. L., Gale, J., & Gander, P. H. (2005). Sleep measurement in flight crew: comparing actigraphic and subjective estimates to polysomnography. *Aviation, Space, and Environmental Medicine*, *76(11)*, 1058-1063.
[23] Littner, M., Kushida, C. A., Anderson, W. M., Bailey, D., Berry, R. B., Davila, D. G., Hirshkowitz, M. et al. (2003). Practice parameters for the role of actigraphy in the study of sleep and circadian rhythms: an update for 2002. *Sleep*, *26(3)*, 337.
[24] Signal, T. L., Gale, J., & Gander, P. H. (2005). Sleep measurement in flight crew: comparing actigraphic and subjective estimates to polysomnography. *Aviation, Space, and Environmental Medicine*, *76(11)*, 1058-1063.

den Berg et al.[25] recently examined whether subjective measurements of in-flight sleep could be a reliable alternative to actigraphic measurements for monitoring pilot fatigue in a large-scale study. Their findings suggested that self-reported sleep duration is a reliable alternative to actigraphic sleep.

Performance can be measured in the field with both embedded and added measures. Embedded measures are part of operational routines and are used to improve aviation safety. Added measures are extrinsic to the operation and typically used for research. An example of an embedded measure is operational flight data monitoring (OFDM). OFDM records flight data from routine operations and analyses these flight data to identify trends and fully investigate the circumstances behind events flagged. Another example of an embedded measure reflects the mandatory/voluntary occurrences reporting (M/VOR, or ASR) facilitating the collection and exchange of information on actual or potential safety hazards and deficiencies and contribute to the prevention of aircraft accidents. The feedback that we received from the 'lessons learned' interviews stressed that although measures such as OFDM and M/VOR could potentially be of high relevance in studies such as ours, it appears that it usually is a big struggle (1) to get the data from the airlines (also a struggle for airlines to gather the data) and (2) to handle the data correctly, synchronizing them to the in-flight measured data points. The literature search showed that these measures were only used in a few studies; with the most recent being from 2012[26].

A good example of an added performance measure is the PVT. This metric is a widely-used test in laboratory settings and in the assessment of performance in real-world activities[27] (including aviation as illustrated by the high frequency of PVT use in the literature review). Cognitive effectiveness is interpreted as the inverse of fatigue and ranges in score from 0 to 100. The PVT is a sustained-attention, reaction-timed task that measures the speed with which subjects respond to a visual stimulus. Research indicates increased sleep debt or sleep deficit correlates with deteriorated alertness, slower problem-solving, declined psycho-motor skills, and increased rate of false responding[28,29,30]. It has been shown to track time of day (circadian) and time awake (homeostatic) factors underlying fatigue in real-world aviation operations[31].

While PVTs in the laboratory are generally set to ten minutes duration, this can be considered a task too long to complete in the field. At least, this is what came out of the 'lessons learned' interviews. Practically, a 10-minute PVT may be considered too intrusive and completed less often than a 5-minute PVT in the field, thus resulting in

[25] Berg, M. J. van den, Wu, L. .J., & Gander, P. H. (2016). Subjective measurements of in-flight sleep, circadian variation, and their relationship with fatigue. *Aviation, Space, and Environmental Medicine*, *87(10)*, 869-875.

[26] Cabon, P., Deharvengt, S., Grau, J. Y., Maille, N., Berechet, I., & Mollard, R. (2012). Research and guidelines for implementing Fatigue Risk Management Systems for the French regional airlines. *Accident Analysis & Prevention*, *45*, 41-44.

[27] Loh, S., Lamond, N., Dorrian, J., Roach, G., & Dawson, D. (2004). The validity of psychomotor vigilance tasks of less than 10-minute duration. *Behavior Research Methods, Instruments, & Computers*, *36(2)*, 339-346.

[28] Dorrian, J., Rogers, N. L., & Dinges, D. F. (2005). Psychomotor vigilance performance: neurocognitive assay sensitive to sleep loss. In: Kushida, C. A. (Ed.), Sleep Deprivation: Clinical Issues, Pharmacology, and Sleep Loss Effects. Marcel Dekker, New York, pp. 39-70.

[29] Balkin, T. J., Bliese, P. D., Belenky, G., Sing, H., Thorne, D. R., Thomas, M., Redmond, D. P., Russo, M., & Wesensten, N. J. (2004). Comparative utility of instruments for monitoring sleepiness-related performance decrements in the operational environment. *Journal of Sleep Research*, *13*, 219-227.

[30] Thorne, D. R., Johnson, D. E., Redmond, D. P., Sing, H. C., Belenky, G., & Shapiro, J. M. (2005). The Walter Reed palm-held psychomotor vigilance test. *Behavior Research Methods*, *37(1)*, 111.

[31] Gander, P. H., Mulrine, H. M., van den Berg, M. J., Smith, A. A., Signal, T. L., Wu, L. J., & Belenky, G. (2015). Effects of sleep/wake history and circadian phase on proposed pilot fatigue safety performance indicators. *Journal of Sleep Research*, *24*, 110-119.

an overall reduction in data collected. The 5-minute PVT is less sensitive to sleepiness than the 10-minute PVT on some PVT-derived metrics. This is especially true for lapses (reaction times greater than 500 milliseconds that are generally thought of as representing a lapse in attention). Mean reaction time, while a sensitive measure of sleepiness in both the 5-minute and 10-minute PVT, is significantly affected by time on task such that mean reaction time increases more during the 10-minute PVT after sleep deprivation relative to the 5-minute PVT[32]. A study designed to validate the 5-minute hand-held PVT showed that there was no significant difference in standardized reaction times between the 5-minute hand-held PVT and 10-minute laboratory PVT[33]. The most apparent advantage of the hand-held PVT is that it is portable and can be used in the workplace.

The literature review showed that in Europe only studies using PDAs were performing in-flight assessments; no paper-based studies were found in Europe assessing aircrew in-flight. This was confirmed by the feedback from the 'lessons learned' interviews saying that, especially in large-scale data collection campaigns, the use of digital devices such as a tablet or smart phone is highly recommended. Data collection applications (or apps) are available that significantly improve the campaigns' efficiency as data handling is digitized completely, making the error-prone task of filling in the output of questionnaire/rating scales into data spreadsheet redundant.

Confidentiality of data

Regarding the protection of gathered data confidentiality, the interviews indicated the need to build strong collaborative relationships with airlines, airlines' safety departments, unions, and regulators. These relationships aid in reducing concerns regarding confidentiality.

Sleep/wake history and performance data is sensitive in the context of workplace performance evaluation and accident analysis and prevention. Researchers, employers, and participants often have concerns over confidentiality of sleep and performance data.

Interviews also revealed the necessity to explain thoroughly that the confidentiality of the data will be guaranteed. A number of interviewees said that they have worked with confidentiality agreements with the airlines. Though relatively time consuming, these are sometimes necessary.

[32] Loh, S., Lamond, N., Dorrian, J., Roach, G., & Dawson, D. (2004). The validity of psychomotor vigilance tasks of less than 10-minute duration. *Behavior Research Methods, Instruments, & Computers*, *36(2)*, 339-346.

[33] Lamond N., Dawson, D., & Roach, G. D. (2005). Fatigue assessment in the field: validation of a hand-held electronic psychomotor vigilance task. *Aviation, Space, and Environmental Medicine*, *76*, 486-489.

Chapter 3: Ranking of aircrew duty periods

This section describes the classification of the six crew duty periods of interest based on the expected level of fatigue. Two different bio-mathematical models were used to estimate the level of fatigue. In addition, a survey was used to provide for a subjective ranking of the duty periods of interest. The following six crew duty periods were considered:

1. Duties of more than 13 hours at the most favourable time of the day;
2. Duties of more than 10 hours at the less favourable time of the day;
3. Duties of more than 11 hours for crew members in an unknown state of acclimatisation;
4. Duties including a high level of sectors (more than six);
5. On-call duties such as standby or reserve, followed by flight duties; and
6. Disruptive schedules.

Bio-mathematical modelling

Each of the six duty periods actually describes a range of possible specific schedules. To calculate fatigue levels with bio-mathematical models, we further redefined the definitions and determined a example duties. Consultation with EASA provided the following clarifications:

- 'Most favourable time of day' is intended to refer to daytime operations (i.e., between 08:00h and 21:59h);
- 'Less favourable time of day' is intended to refer to operations that encroach (part of) the night (i.e., the period between 02:00h and 04:59h); and
- 'Disruptive schedule' refers to repetitive early starts, late finishes, night duties, and combinations thereof.

This information was then used to create schedules that could be fed into the bio-mathematical models. The schedules were selected such that they represent realistic (although not necessarily common) flight duties.

For duty type 3 (where crew members are in an unknown state of acclimatisation), it was assumed that the time difference between reference time and local time where the crew starts the next duty is 12 hours, and the time elapsed since reporting at reference time is 48 hours (as specified in Annex II to Regulation 965/2012).

For duty type 4 (high level of sectors), the maximum possible number of sectors 10 was assumed (as specified in Annex II to Regulation 965/2012).

Duty type 6 (disruptive schedules) was assumed to involve four consecutive early starts or four consecutive late finishes.

All times are in reference time (the local time at the reporting point situated in a 2-hour wide time zone band around the local time where a crew member is acclimatised). The assumption was that crews are acclimatised to the local time of the departure time zone, except for flight duty period number 3.

1. Duties of more than 13 hours at the most favourable time of the day. This refers to daytime operations (from 08:00h to 21:59h);
   - Duty started at 08:00h and ended at 20:59h;
2. Duties of more than 10 hours at the less favourable time of the day. This refers to operations that encroach (part of) the night (the period between 02:00h and 04:59h);
   - Duty started at 19:00h and ended next day at 05:59h;

3. Duties of more than 11 hours for crew members in an unknown state of acclimatisation;
   - Crew arrived at 11:00h with a time zone difference +12. Duty started 48 hours later at 11:00h and ends 11 hours later at 21:59h;
4. Duties including a high level of sectors (more than six). This refers to daytime operations (from 08:00h to 21:59h);
   - Duty started at 08:00h and ended at 16:59h. Duty included 10 sectors;
5. On-call duties such as standby or reserve followed by flight duties. This refers to daytime flight duties;
   - Standby started at 06:00h. Duty started at 11:00h and ended at 23:59h.
6. Disruptive schedules. This refers to repetitive early starts, late finishes, night duties, and combinations thereof;
   a. Early starts: Four consecutive flight duties starting at 05:00h and ending at 14:59h;
   b. Late finishes: Four consecutive flight duties starting at 16:00h and ending next day at 01:59h; and
   c. Night duties: Four consecutive flight duties starting at 23:00h and ending next day at 08:59h.

Furthermore, for each specific flight duty period, fatigue levels were calculated from two initial conditions: (1) Crew is fully rested at the start of the duty period and (2) crew is pre-fatigued at the start of the duty period. For fully rested crew, the last sleep episode (duration of eight hours) ends two hours before start of duty unless stated otherwise. For pre-fatigued crew, the first three hours of the predicted final sleep episode before the start of duty was regarded as 'awake', resulting in a sleep duration of five hours. Additionally, the following assumptions were made:

- The window of circadian low (WOCL) ranges from 02:00h to 05:59h in the time zone to which a crew member was acclimatised;
- Transfer time from bed to start of the flight duty or from the end of the flight duty to bed was two hours;
- Sleep was assumed to not occur after the start of on-call duty period;
- Application of fatigue risk management (FRM), in-flight rests, and/or augmented flight crew were excluded;
- Extensions and commander's discretion were not used;
- A flight duty included two sectors unless stated otherwise; and
- Quality of sleep was scored with the highest value.

The duties were analysed by using Boeing Alertness Model (BAM, CrewAlert Pro 3.9.7) and Sleep, Activity, Fatigue, and Task Effectiveness, Fatigue Avoidance Scheduling Tool (SAFTE-FAST, v1.2.4.92). BAM predicts alertness on common alertness scale (CAS) from 0 (least alert state) to 10,000 (most alert state). CAS is linearly mapped against the KSS[34] where a KSS value of 9 (very sleepy, great effort to keep awake, fighting sleep) maps to 0 CAS points and KSS of 1 (extremely alert) maps to 10,000 CAS points. The output of the SAFTE model provides a percentage of performance effectiveness (Effect) from 0 (low effectiveness) to 100 (high effectiveness).

[34] KSS is a 9-point scale: 1. Extremely alert, 2. Very Alert, 3. Alert, 4. Rather alert, 5. Neither alert nor sleepy, 6. Some signs of sleepiness, 7. Sleepy, but no difficulty remaining awake, 8. Sleepy, some effort to keep alert, 9. Very sleepy, great effort to keep awake, fighting sleep.

Survey

A survey was used to provide for a subjective ranking of the duty periods of interest based on the associated fatigue level. In addition, a survey was used to identify potential fatigue hotspots in these flight duty periods.

The survey ranking inputs were presented and discussed in an Addendum to D1 (Definition of the Baseline).

The survey was developed in a number of iterations in order to ensure high-quality questions using a language and format that is easy to understand for the participants. First, the survey outline was designed and a concept survey was developed in a text formatting program for ease of adjustability. This first set-up of the survey was then reviewed by a small committee within the consortium completed by two commercial pilots working for NLR. The results of this review were discussed and adjustments were made in the next iteration. This next set-up of the survey was reviewed and commented on again by the same committee. After two iterations, the survey was transferred to LimeSurvey to include the sequencing of the questions and to also define the format of the data output. A final review was performed by a group of 20 participants, including commercial pilots and cabin crew, and the full project consortium.

The survey questions asked aircrew respondents to assign a fatigue rating (using the KSS) for each of the six flight duty periods they have experienced in the past three years. From the ratings, a rank ordering could then be derived.

The survey could be accessed with any type of computer or mobile device with internet access. Data is to be gathered via LimeSurvey which is a software package for surveys. Using LimeSurvey, the data gathered can be saved on a server within the consortium making the data available for the consortium, but not for third parties. The package is relatively easy to use, but hardly limits possibilities for adjustments. Furthermore, the data can be directly imported to most data analysis software packages. The package allows for anonymous answering.

The fact that the survey was anonymized allows for filling out the survey multiple times. This was countered by using cookies to make it more difficult to fill in the survey twice and by gathering IP addresses to be able to inspect if the same address was used multiple times. If analysis showed that the same IP address was used more than once, the corresponding survey output was checked for similarities. In case of high resemblance of the outputs, only a single stream of survey output was used in the data analysis. The IP addresses were used only for the purpose of this inspection and were removed from the dataset after this inspection.

The data analysis started with analysing if the gathered sample represents the selected population. This means that the group of participants represented the complete population of aircrew on several grounds, such as gender, country of home base, passengers versus cargo, etc. The data was split per group (i.e., pilots, cabin crew, researchers, and safety personnel) and for each group counts were done for the various questions and ratings on the duty periods were compared to create an overall input for the ranking for all aircrew duty types. Finally, the demographic data was explored with the ranking to explore possible relationships.

Classification of aircrew duty periods per source

Output bio-mathematical modelling

Results for BAM reflect the CAS values for the analysed FDP; the lower the values, the lower the level of alertness. Results for the SAFTE model reflect the lowest value of the effectiveness score; the lower the values, the lower level of effectiveness. See Table 2 for the outputs of the bio-mathematical modelling; the ranking indicated by the model calculations is shown in brackets (1 representing the highest ranking and 8 the lowest).

Table 2 Results from FDP calculations with bio-mathematical models

Flight duty type	***BAM***		***SAFTE model***	
	Rested	Pre-fatigued	Rested	Pre-fatigued
1) > 13 hours at favourable time	(7) 4709	(7) 4156	(7/8) 95	(7/8) 90
2) > 10 hours at unfavourable time	(1) 2403	(1) 1436	(2) 70	(2) 64
3) > 11 hours for unknown state	(5) 3560	(5) 2204	(6) 89	(5) 80
4) High number of sectors	(8) 5738	(8) 5090	(7/8) 95	(7/8) 90
5) On call duties	(2) 2587	(4) 2103	(5) 83	(6) 85
6 a) Cumulative early starts	(6) 4502	(6) 3315	(4) 80	(4) 78
6 b) Cumulative late finishes	(3) 3095	(2) 1905	(3) 79	(3) 70
6 c) Cumulative night duties	(4) 3132	(3) 1946	(1) 66	(1) 56

Many of the available bio-mathematical models are fundamentally based on the two-process model of sleep regulation that describes the interaction between a homeostatic process and a circadian process[35]. Even when models have a similar basis, different values may be incorporated for the parameters that are used in the equations for describing the homeostatic and circadian process. Examples are the parameters for the rate at which fatigue increases with time awake and the parameter that describes the amplitude of the circadian process. Differences in results of various models (using the same input) may be explained by different choices for such parameter values.

When considering the results from Table 2, the following limitations of modelling should be taken into account:

- The contribution of multiple sectors in BAM is tuned to collected data, but is relatively small compared to the main components: time awake, time of day, and prior sleep debt;
- The SAFTE model effectiveness score does not consider sectors;
- Model outputs represent the population average and may not be accurate for specific individuals;
- Model sleep predictions may not reflect actual sleep which is fundamental to the validity of its output; and
- Models may not take into account the operational context and mitigations.

[35] Mallis, M. M., Meijdal, S., Nguyen, T. T., & Dinges, D. F. (2004). Summary of the key features of seven biomethematical models of human fatigue and performance. *Aviation, Space and Environmental Medicine, 75(3)*, A4-14.

Output survey

As the results of the survey are pending, the ranking of the flight duty periods cannot be finalized in D1 (Definition of the Baseline).

The two duty periods ranked as the most fatiguing

This section is to be completed in an Addendum to D1 (Definition of the Baseline).

The ranking results were based on the findings from the bio-mathematical models and survey and were corroborated (or rejected) by scientific findings. No interim conclusions are drawn yet. The following recent publications provide good examples of relevant insights to support the ranking of aircrew duty periods.

In a study by Honn et al.[36], regional airline pilots flew two duty days in a high-fidelity, moving-base, full-flight, regional jet flight simulator. One of the two duty days involved flying five segments, as is typical for regional airline operations; the other duty day involved flying a single (longer) segment, while the duty duration (nine hours) and start time (05:15h) were the same. During each of the two duty days, fatigue test bouts (PVT, SP and KSS) were administered ten times. Both objective measurements and subjective reports revealed greater build-up of fatigue in the five-segment duty day than in the single-segment duty day. The experiment was designed such that the additional fatigue in the multi-segment duty day could be attributed specifically to the multiple flight segments, and the associated multiple take-offs and landings. Other factors that could systematically affect fatigue were purposely standardized and controlled for. As such, the study findings imply that the task load associated with multiple take-offs and landings in multi-segment operations results in increased fatigue over the duty day. This is consistent with data from laboratory-based fatigue studies in which task load was manipulated[37,38,39].

Sallinen et al.[40] conducted a field study on a representative sample of the airline pilots of a medium-sized airline. The sample consisted of 90 pilots, of whom 30 flew long-haul routes, 30 short-haul routes, and 30 flew both. A total of 86 pilots completed the measurements that lasted for almost two months per pilot. Results showed that short- and long-haul duty periods covering the whole domicile night (00:00h – 05:59h at home base) were most consistently associated with reduced sleep-wake ratio and subjective alertness. The results also showed that the pilots tended to increase the use of effective on-duty alertness management strategies (consuming alertness-promoting products and taking strategic naps) in connection with the flight duty periods that overlapped the domicile night. The results suggest that flight duty periods covering the whole domicile night should be prioritised over the other flight duty periods in fatigue management, regardless of whether a flight duty period is a short- or a long-haul. The finding of reduced sleep sufficiency and subjective alertness

---

[36] Honn, K. A., Satterfield, B. C., McCauleya, P., Caldwell, J. L., & van Dongen, H. P. A. (2016). Fatiguing effect of multiple take-offs and landings in regional airline operations. *Accident Analysis & Prevention*, *86*, 199–208.

[37] Dongen, H. P. A. van, & Dinges, D. F. (2007). Individual differences in response to sleep deprivation. Final Technical Report, NIH R01 HL070154.

[38] Lim, J., Wu, W. C., Wang, J., Detre, J. A., Dinges, D. F., & Rao, H. (2010). Imaging brainfatigue from sustained mental workload: an ASL perfusion study of the time-on-task effect. *NeuroImage*, *49*, 3426–3435.

[39] Goel, N., Abe, T., Braun, M. E., & Dinges, D. F. (2014). Cognitive workload and sleep restriction interact to influence sleep homeostatic responses. *Sleep*, *37*, 1745-1756.

[40] Sallinen, M., Sihvolaa, M., Puttonena, S., Ketolac, K., Tuoric, A., Härmäa, M., Kecklundd, G., & Åkerstedt, T. (2017). Sleep, alertness and alertness management among commercial airline pilots on short-haul and long-haul flights. *Accident Analysis & Prevention*, *98*, 320-329.

especially in connection with whole night flight duty periods is well in line with a number of previous studies conducted on airline pilots and other groups of transport professionals[41,42,43,44]. In addition to the whole night flight duty period, the short- and long-haul flight duty periods that covered either the first or second part of the night were, to some extent, associated with a lowered sleep-wake ratio and subjective alertness. This finding is in accordance with previous studies conducted on a wide range of occupational groups[45,46,47].

In their research, Vejvoda et al.[48] demonstrated that short-haul pilots experienced moderate to severe fatigue when finishing flight duty periods late at night. These fatigue levels exceeded those observed after duty periods with early starts, despite the fact that duty period duration was shorter and prior sleep period time was longer. Pilots on late-finishing FDPs (i.e., duty start after 17:00h) were awake longer by an average of 5.5 hour (6.6 versus 1.1 hour) before commencing their duty than pilots who started early in the morning. Late-finishing flights were associated with long times awake at a time when the circadian system stops promoting alertness, and an increased, previously underestimated fatigue risk. Other studies have identified sleep duration, time of day, number of flights and duty duration to influence pilots' fatigue[49,50]. The results in short-haul pilots indicate that the times spent awake not only during but also prior to a FDP are strong predictors of fatigue at duty end, apparently outweighing the effects of the observed variation in the prior sleep duration.

[41] Eriksen, C. A., & Åkerstedt, T. (2006). Aircrew fatigue in trans-Atlantic morning and evening flights. *Chronobiology International*, *23(4)*, 843-858.
[42] Gander, P. H., Mulrine, H. M., van den Berg, M. J., Smith, A. A., Signal, T. L., Wu, L. J., & Belenky, G. (2015). Effects of sleep/wake history and circadian phase on proposed pilot fatigue safety performance indicators. *Journal of Sleep Research*, *24(1)*, 110-119,
[43] Härmä, M., Sallinen, M., Ranta, R., Mutanen, P., & Müller, K. (2002). The effect of an irregular shift system on sleepiness at work in train drivers and railway traffic controllers. *Journal of Sleep Research*, *11*, 141-151.
[44] Sallinen, M., & Kecklund, G. (2010). Shift work, sleep, and sleepiness – differences between shift schedules and systems. *Scandinavian Journal of Work, Environment & Health*, *36(2)*, 121-133.
[45] Sallinen, M., & Hublin, C. (2015). Fatigue-inducing factors in transportation operations. In: Stephen Popkin, M. (Ed.), Reviews of Human Factors and Ergonomics: Worker Fatigue and Transportation Safety, vol. 10. SAGE Publications, pp. 138-173.
[46] Pylkkönen, M., Sihvola, M., Hyvärinen, H. K., Puttonen, S., Hublin, C., & Sallinen, M. (2015). Sleepiness, sleep, and use of sleepiness countermeasures in shift-working long-haul truck drivers. *Accident Analysis & Prevention*. *80*, 201-210.
[47] Roach, G.D., Sargent, C., Darwent, D., & Dawson, D. (2012). Duty periods with early start times restrict the amount of sleep obtained by short-haul airline pilots. *Accident Analysis & Prevention*, *45*, 22-26.
[48] Vejvoda, M., Elmenhorst, E.M., Pennig, S.B., Parh, G., Maass, H., Tritschler, K., Basner, M., & Aeschbach, D. (2014). Significance of time awake for predicting pilots' fatigue on short-haul flights: implications for flight duty time regulations. *Journal of Sleep Research*, *23(5)*, 564-567.
[49] Powell, D. M., Spencer, M. B., Holland, D., Broadbent, E. & Petrie, K. J. (2007). Pilot fatigue in short-haul operations: effects of number of sectors, duty length, and time of day. *Aviation, Space, and Environmental Medicine*, *78*, 698-701.
[50] Petrilli, R. M., Roach, G. D., Dawson, D., & Lamond, N. (2006). The sleep, subjective fatigue, and sustained attention of commercial airline pilots during an international pattern. *Chronobiology International*, *23*, 1357-1362.

Chapter 4: Characterisation of the selected population

For purposes of data collection, a representative population of Member States, air transport operators and type of operations has been identified.

Selecting a balanced set of EU air operators and operations

The identification of a representative population included the following steps:

- Establishment of a subset of Member States representative of conditions in the EU aviation sector as a whole. The country grouping shall encompass the type of air transport operators and related envelope of operations that are typical for the EU aviation market; and
- Definition of criteria for the conduct of a screening of CAT aeroplane operators to achieve a representative mapping of air operations considering the following set of characteristics:

Table 3 Screening criteria of CAT operators

Type	*Flight duration*
Long-haul	More than 5 hours and crossing 3 time zones
Medium-haul	More than 2 hours
Short-haul	Between 1 and 2 hours' duration
Regional	Less than 1 hour
Sole cargo flights	-

The resulting representative set of air operators and operations is to be used in the context of subsequent data gathering activities.

Approach EU aviation ensemble

An overview of European CAT operators classified by EASA Member State and sub-classified by type of operation was assembled. Four geographical regions were defined covering Europe: East, West, North, and South (see Table 4). Regarding type of operation, an internet search was performed and expert opinions were gathered to determine the air operators per Member States.

Table 4 Geographical regions within Europe

Region 1	*Region 2*	*Region 3*	*Region 4*
North Europe	*West Europe*	*South Europe*	*East Europe*
Denmark	United Kingdom	Italy	Romania
Sweden	Germany	Spain	Slovakia
Norway	Netherlands	Greece	Czech Republic
Finland	Ireland	Cyprus	Bulgaria
Iceland	Austria	Malta	Poland
	Belgium	Portugal	Hungary
	Liechtenstein	Croatia	Estonia
	Luxembourg	Slovenia	Lithuania
	Switzerland		Latvia
	France		

Note that the size of air operations (based on number of aircraft) is the largest within Region 2 West Europe (around 3300 aircraft operational), followed by Region 2 South Europe (with around 850 aircraft) and Region 1 North Europe (with around 550 aircraft) and finally Region 4 East Europe (with around 360 aircraft).

Approach screening CAT

Criteria were defined to be able to narrow the EU aviation ensemble that was determined. The following criteria were taken into account in the screening of CAT operator:

- Volume of air operations (as a function of number of aircraft) in various types of operation. This was to provide insight in the (potential) exposure to fatigue. An internet search was performed to determine this; these numbers were gathered and averaged based on the multiple information sources available on the internet;
- The extent to which operators use deviations/derogations from the EU FTL Regulations. This was based on information on deviations and derogations from the EASA website; and
- The type of FDPs (i.e., which of the six duty periods of interest) that are operated by the operators. The operators should operate the two duty periods ranked as the most fatiguing for inclusion to be possible. This particular criterion can only be applied after the ranking of aircrew duty periods is completed.

Selected population

These criteria resulted in a set of EU air operators and operations (preliminary as the criterion of operating the two duty periods ranked as the most fatiguing is still to be applied).

The two largest (in number of aircraft) of each suitable operator in the regions are presented in selection 1. Numbers three and four are in selection 2 (i.e., the operators to target second) and numbers five and six in selection 3 (i.e., the operators to target third). In selection 3 both Region North and South show three candidate operators; i.e., a sole regional operator was added for both regions. All selections contain operations from short-, medium-, long-haul, and regional.

The candidate airline operators selected were asked to participate in the planned data gathering activities. Operators in selection 1 were approached first; hereafter operators in selections 2, 3 and 4; i.e., depending on the willingness of the operators to participate and the number of aircrew volunteering per operator[51].

Table 5 Airline operator selection 1

Selection 1 CAT	*Region*	*Long*	*Medium*	*Short*	*Regional*
WIZZ Air	East		X	X	
LOT Polish Airlines	East	X	X	X	X
Lufthansa	West	X	X	X	X
Ryanair	West		X	X	
Scandinavian Airlines	North	X	X	X	X
Norwegian Air Shuttle	North		X	X	
Vueling	South		X	X	
Alitalia	South	X	X	X	

[51] In Deliverable 2.2 (Definition of Data Collection Process) the required sample size of aircrew for the data collection campaign was calculated.

Table 6 Airline operator selection 2

Selection 2 CAT	***Region***	***Long***	***Medium***	***Short***	***Regional***
Air Baltic	East		X	x	
TAROM	East		X	X	X
British Airways	West	X	X	X	
KLM	West	X	X	X	
Norwegian Air Int.	North	X	X	X	
Icelandair	North	X	X	X	
Iberia	South	X	X	X	
TAP Portugal	South	X	X	X	

Table 7 Airline operator selection 3

Selection 3 CAT	***Region***	***Long***	***Medium***	***Short***	***Regional***
Czech Airlines	East	X	X	X	
Smartwings	East		X	X	
Air Berlin	West	X	X	X	X
Flybe	West			X	X
BRA Braathens	North				X
Thomas Cook Scan.	North	X	X	X	
WOW Air	North	X	X	X	
Air Europa	South	X	X	X	
Aegean Airlines	South		X	X	
Air Nostrum	South				X

Note that Smartwings, Air Berlin, Thomas Cook Scandinavia, Air Europa, and Aegean Airlines are the only (full or part) leisure operators in the three selections.

Regarding sole cargo, the following operators would be suitable candidates: ASL Airlines Belgium, Cargolux and Lufthansa Cargo (all Region West).

As the ranking of the flight duty periods cannot be finalized in D1, the final selected population was completed in an Addendum to D1 (Definition of the Baseline).

References

Åkerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience*, *52*, 29-37.

Åkerstedt, T., Anund, A., Axelsson, J., & Kecklund, G. (2014). Subjective sleepiness is a sensitive indicator of insufficient sleep and impaired waking function. *Journal of Sleep Research*, *23*, 240-252.

Balkin, T. J., Bliese, P. D., Belenky, G., Sing, H., Thorne, D. R., Thomas, M., Redmond, D. P., Russo, M., & Wesensten, N. J. (2004). Comparative utility of instruments for monitoring sleepiness-related performance decrements in the operational environment. *Journal of Sleep Research*, *13*, 219-227.

Berg, M. J. van den, Signal T. L., Mulrine, H. M., Smith, A. A. T., Gander, P. H., & Serfontein, W. (2015). Monitoring and managing cabin crew sleep and fatigue during an ultra-long range trip. *Aerospace Medicine and Human Performance*, *86(8)*, 705-713.

Berg, M. J. van den, Signal, T. L., Mulrine, H., Smith, A. A. T., & Gander, P. H. (2013). Evaluation of the sleep and performance of South African Airways cabin crew on the Johannesburg-New York ultra long range flight. Sleep/Wake Research Centre, Massey University.

Berg, M. J. van den, Wu, L. J., & Gander, P. H. (2016). Subjective measurements of in-flight sleep, circadian variation, and their relationship with fatigue. *Aviation, Space, and Environmental Medicine*, *87(10)*, 869-875.

Brown, L., Schoutens, A. M. C., Whitehurst, G., Booker, T. J., Davis, T., Losinski, S., & Diehl, R. (2014). The effect of blue light therapy on flight crew-members behavioral alertness. Available at SSRN: https://ssrn.com/abstract=2402409.

Cabon, P., Deharvengt, S., Grau, J. Y., Maille, N., Berechet, I., & Mollard, R. (2012). Research and guidelines for implementing Fatigue Risk Management Systems for the French regional airlines. *Accident Analysis & Prevention*, *45*, 41-44.

Dongen, H. P. A. van, & Dinges, D. F. (2007). Individual differences in response to sleep deprivation. Final Technical Report, NIH R01 HL070154.

Dorrian, J., Rogers, N. L., & Dinges, D. F. (2005). Psychomotor vigilance performance: neurocognitive assay sensitive to sleep loss. In: Kushida, C. A. (Ed.), Sleep Deprivation: Clinical Issues, Pharmacology, and Sleep Loss Effects. Marcel Dekker, New York, pp. 39-70.

Eriksen, C. A., & Åkerstedt, T. (2006). Aircrew fatigue in trans-Atlantic morning and evening flights. *Chronobiology International*, *23(4)*, 843-858.

Ferguson, S. A., Paech, G. M., Sargent, C., Darwent, D., Kennaway, D. J., & Roach, G. D. (2012). The influence of circadian time and sleep dose on subjective fatigue ratings. *Accident Analysis & Prevention*, *45*, 50-54.

Gander, P. H., Mangie, J., van den Berg, M. J., Smith, A. A. T., Mulrine, H. M., & Signal, T. L. (2014). Crew fatigue safety performance indicators for fatigue risk management systems. *Aviation, Space, and Environmental Medicine*, *85(2)*, 139-147.

Gander, P. H., Mulrine, H. M., van den Berg, M. J., Smith, A. A. T., Signal, T. L., Wu, L. J., & Belenky, G. (2014). Pilot fatigue: relationships with departure and arrival times, flight duration, and direction. *Aviation, Space, and Environmental Medicine*, *85(8)*, 833-840.

Gander, P. H., Mulrine, H. M., van den Berg, M. J., Smith, A. A., Signal, T. L., Wu, L. J., & Belenky, G. (2015). Effects of sleep/wake history and circadian phase on proposed pilot fatigue safety performance indicators. *Journal of Sleep Research*, *24(1)*, 110-119.

Gander, P. H., Signal, T. L., van den Berg, M. J., Mulrine, H. M., Jay, S. M., & Mangie, J. C. (2013). In-flight sleep, pilot fatigue and Psychomotor Vigilance Task performance on ultra-long range versus long range flights. *Journal of Sleep Research*, *22(6)*, 697-706.

Gander, P. H., van den Berg, M. J., Mulrine, H. M., & Mangie, J. (2013). Circadian adaptation of airline pilots during extended duration operations between the USA and Asia. *Chronobiology International*, *30(8)*, 963-972.

Gander, P., van den Berg, M., Jay, S., & Signal, T. (2012). Comparison of flight crew sleep and fatigue during Delta Air Lines long range and ultra-long range operations. Sleep/Wake Research Centre, Massey University.

Gillberg, M., Kecklund, G., & Åkerstedt, T. (1994). Relations between performance and subjective ratings of sleepiness during a night awake. *Sleep*, *17(3)*, 236-241.

Goel, N., Abe, T., Braun, M. E., & Dinges, D. F. (2014). Cognitive workload and sleep restriction interact to influence sleep homeostatic responses. *Sleep*, *37*, 1745-1756.

Greeley, H. P., Roma, P. G., Mallis, M. M., Hursh, S. R., Mead, A. M., & Nesthus, T. E. (2013). Field study evaluation of cepstrum coefficient speech analysis for fatigue in aviation cabin crew. FAA Report No. DOT/FAA/AM-13/19.

Härmä, M., Sallinen, M., Ranta, R., Mutanen, P., & Müller, K. (2002). The effect of an irregular shift system on sleepiness at work in train drivers and railway traffic controllers. *Journal of Sleep Research*, *11*, 141-151.

Holmes, A., Al-Bayat, S., Hilditch, C., & Bourgeois-Bougrine, S. (2012). Sleep and sleepiness during an ultra long-range flight operation between the Middle East and United States. *Accident Analysis & Prevention*, *45*, 27-31.

Honn, K. A., Satterfield, B. C., McCauleya, P., Caldwell, J. L., & van Dongen, H. P. A. (2016). Fatiguing effect of multiple take-offs and landings in regional airline operations. *Accident Analysis & Prevention*, *86*, 199-208.

Houston, S., Dawson, K., & Butler, S. (2012). Fatigue reporting among aircrew: incidence rate and primary causes. *Aviation, Space, and Environmental Medicine*, *83*, 800-804.

Ingre, M., van Leeuwen, W., Klemets, T. et al*.* (2014). Validating and extending the three process model of alertness in airline operations. *PLoS One*, 9.

Kaida, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y. et al. (2006). Validation of the Karolinska sleepiness scale against performance and EEG variables. *Clinical Neurophysiology*, *117(7)*, 1574-1581.

Lamond N., Dawson, D., & Roach, G. D. (2005). Fatigue assessment in the field: validation of a hand-held electronic psychomotor vigilance task. *Aviation, Space, and Environmental Medicine*, *76*, 486-489.

Lim, J., Wu, W. C., Wang, J., Detre, J. A., Dinges, D. F., & Rao, H. (2010). Imaging brainfatigue from sustained mental workload: an ASL perfusion study of the time-on-task effect. *NeuroImage*, *49*, 3426-3435.

Littner, M., Kushida, C. A., Anderson, W. M., Bailey, D., Berry, R. B., Davila, D. G., Hirshkowitz, M. et al. (2003). Practice parameters for the role of actigraphy in the study of sleep and circadian rhythms: an update for 2002. *Sleep*, *26(3)*, 337.

Loh, S., Lamond, N., Dorrian, J., Roach, G., & Dawson, D. (2004). The validity of psychomotor vigilance tasks of less than 10-minute duration. *Behavior Research Methods, Instruments, & Computers*, *36(2)*, 339-346.

Mallis, M. M., Meijdal, S., Nguyen, T. T., & Dinges, D.F. (2004). Summary of the key features of seven biomethematical models of human fatigue and performance. *Aviation, Space and Environmental Medicine*, *75(3)*, A4-14.

Mello, M. T. de, Esteves, A. M., Pires, M. L. N., Santos, D. C., Bittencourt, L. R. A., Silva, R. S., & Tufik, S. (2008). Relationship between Brazilian airline pilot errors and time of day. *Brazilian Journal of Medical and Biological Research*, *41(12)*, 1129-1131.

O'Hagan, A. D., Issartel, J., Fletcher, R., & Warrington, G. (2016). Duty hours and incidents in flight among commercial airline pilots. *International Journal of Occupational Safety and Ergonomics*, *22(2)*, 165-172.

Pascoe, P. A., Johnson, M. K., Roberston, K. A., & Spencer, M. B. (1995). Sleep in rest facilities on board aircraft: field studies. DERA Report No. DERA/CHS/A&N/CR/95/002.

Petrilli, R. M., Roach, G. D., Dawson, D., & Lamond, N. (2006). The sleep, subjective fatigue, and sustained attention of commercial airline pilots during an international pattern. *Chronobiology International*, *23*, 1357-1362.

Powell, D. M. C., Spencer, M. B., & Petrie, K. J. (2010). Fatigue in airline pilots after an additional day's layover period. *Aviation, Space, and Environmental Medicine*, *81*, 1013-1017.

Powell, D. M., Spencer, M. B., & Petrie, K. J. (2011). Automated collection of fatigue ratings at the top of descent: a practical commercial airline tool. *Aviation, Space, and Environmental Medicine*, *82(11)*, 1037-1041.

Powell, D. M., Spencer, M. B., Holland, D., Broadbent, E. & Petrie, K. J. (2007). Pilot fatigue in short-haul operations: effects of number of sectors, duty length, and time of day. *Aviation, Space, and Environmental Medicine*, *78*, 698-701.

Powell, D., Spencer, M. B., Holland, D., & Petrie, K. J. (2008). Fatigue in two-pilot operations: implications for flight and duty time limitations. *Aviation, Space, and Environmental Medicine*, *79*(11), 1047-1050.
Pylkkönen, M., Sihvola, M., Hyvärinen, H. K., Puttonen, S., Hublin, C., & Sallinen, M. (2015). Sleepiness, sleep, and use of sleepiness countermeasures in shift-working long-haul truck drivers. *Accident Analysis & Prevention*, *80*, 201-210.

Reis, C., Mestre, C., & Canhão, H. (2013). Prevalence of fatigue in a group of airline pilots. *Aviation, Space, and Environmental Medicine*, *84(8)*, 828-833.

Reis, C., Mestre, C., Canhão, H., Gradwell, D., & Paiva, T. (2016). Sleep and fatigue differences in the two most common types of commercial flight operations. *Aerospace Medicine and Human Performance*, *87(9)*, 811-815.

Reis, C., Mestre, C., Canhão, H., Gradwell, D., & Paiva, T. (2016). Sleep complaints and fatigue of airline pilots. *Sleep Science*, *9(2)*, 73-77.

Roach, G. D., Darwent, D., Sletten, T. L., & Dawson, D. (2011). Long-haul pilots use in-flight napping as a countermeasure to fatigue. *Applied Ergonomics*, *42*, 214-218.

Roach, G. D., Petrilli, R. M. A., Dawson, D., & Lamond, N. (2012). Impact of layover length on sleep, subjective fatigue levels, and sustained attention of long-haul airline pilots. *Chronobiology International*, *29(5)*, 580-586.

Roach, G.D., Sargent, C., Darwent, D., & Dawson, D. (2012). Duty periods with early start times restrict the amount of sleep obtained by short-haul airline pilots. *Accident Analysis & Prevention*, *45*, 22-26.

Robertson, K. A., Spencer, M. B., Stone, B. M., & Johnson, M. K. (1997). Scheduling the on-board rest of aircrew. DERA Report No. DERA/CHS/PP5/CR97095/1.0.

Roma, P. G., Hursh, S. R., Mead, A. M., & Nesthus, T. E. (2012). Flight attendant work/rest patterns, alertness, and performance assessment: field validation of biomathematical fatigue modeling. FAA report No. DOT/FAA/AM-12/12.

Roma, P. G., Mallis, M. M., Hursh, S. R., Mead, A. M., & Nesthus, T. E. (2010). Flight attendant fatigue recommendation II: flight attendant work/rest patterns, alertness, and performance assessment. FAA Report No. DOT/FAA/AM-10/22.

Sallinen, M., & Hublin, C. (2015). Fatigue-inducing factors in transportation operations. In: Stephen Popkin, M. (Ed.), Reviews of Human Factors and Ergonomics: Worker Fatigue and Transportation Safety, vol. 10. SAGE Publications, pp. 138-173.

Sallinen, M., & Kecklund, G. (2010). Shift work, sleep, and sleepiness – differences between shift schedules and systems. *Scandinavian Journal of Work, Environment & Health*, *36(2)*, 121-133.

Sallinen, M., Sihvolaa, M., Puttonena, S., Ketolac, K., Tuoric, A., Härmäa, M., Kecklundd, G., & Åkerstedt, T. (2017). Sleep, alertness and alertness management among commercial airline pilots on short-haul and long-haul flights. *Accident Analysis & Prevention*, *98*, 320-329.

Samel, A., Wegmann, H. M., Vejoda, M., Drescher, E. E. J., Gundel, A., Manzey, D., & Wenzel, J. (1997). Two-crew operations: stress and fatigue during long-haul night flights. *Aviation, Space, and Environmental Medicine*, *68(8)*, 679-687.

Samel, A., Wegmann, H.-M., & Vejvoda, M. (1997). Aircrew fatigue in long-haul operations. *Accident Analysis & Prevention*, *29(4)*, 439-452.

Samn, S. W., & Perelli, L. P. (1982). Estimating aircrew fatigue: a technique with implications to airlift operations. USAF School of Aerospace Medicine Technical Report No. SAM-TR-82-21.

Signal, T. L., Gale, J., & Gander, P. H. (2005). Sleep measurement in flight crew: comparing actigraphic and subjective estimates to polysomnography. *Aviation, Space, and Environmental Medicine*, *76(11)*, 1058-1063.

Signal, T. L., Gander, P. H., van den Berg, M.J., & Graeber, R. C. (2013). In-flight sleep of flight crew during a 7-hour rest break: implications for research and flight safety. *Sleep*, *36(1)*, 109-115.

Signal, T. L., Mulrine, H. M., van den Berg, M.J., Smith, A. A. T., Gander, P. H., & Serfontein, W. (2014). Mitigating and monitoring flight crew fatigue on a westward ultra-long-range flight. *Aviation, Space, and Environmental Medicine*, *85*, 1199-1208.

Srivistava, A. S., & Barton, P. (2012). Collaboration on the human factors monitoring program (HFMP) study. NASA Report No. TM-2012-216053.

Thomas, M. J. W., & Ferguson, S. A. (2010). Prior sleep, prior wake, and crew performance during normal flight operations. *Aviation, Space, and Environmental Medicine*, *81(7)*, 665-670.

Thomas, M. J. W., Petrilli, R., M., & Roach, G. D. (2007). The Impacts of Australian Transcontinental 'Back of Clock' Operations on Sleep and Performance in Commercial Aviation Flight Crew. ATSB report.

Thorne, D. R., Johnson, D. E., Redmond, D. P., Sing, H. C., Belenky, G., & Shapiro, J. M. (2005). The Walter Reed palm-held psychomotor vigilance test. *Behavior Research Methods*, *37(1)*, 111.

Vejvoda, M., Elmenhorst, E.M., Pennig, S.B., Parh, G., Maass, H., Tritschler, K., Basner, M., & Aeschbach, D. (2014). Significance of time awake for predicting pilots' fatigue on short-haul flights: implications for flight duty time regulations. *Journal of Sleep Research*, *23(5)*, 564-567.

Wu, L. J. (2013). Evidence based fatigue risk management during 24/7 operations: objective assessment of pilots' sleep, performance, and fatigue during ultra long range and long range flights. Doctoral dissertation, Washington State University.

List of abbreviations

Abbreviations	***Description***
AOC	Air Operator Certificate
ASR	Air Safety Report
BAM	Boeing Alertness Model
CAS	Common Alertness Scale
CAT	Commercial Air Transport
D	Deliverable
EASA	European Aviation Safety Agency
EC	European Commission
EEG	Electroencephalography
ESS	Epworth Sleepiness Scale
EU	European Union
FAST	Fatigue Avoidance Scheduling Tool
FDM	Flight Data Monitoring
FDP	Flight Duty Period
FOQA	Flight Operations Quality Assurance
FRM	Fatigue Risk Management
FRMS	Fatigue Risk Management System
FSS	Fatigue Severity Scale
FTL	Flight Time Limitation
JSS	Jenkins Sleep Scale
KSS	Karolinska Sleepiness Scale
M/VOR	Mandatory/Voluntary Occurrences Reporting
NASA	National Aeronautics and Space Administration
NASA TLX	NASA Task Load Index
OFDM	Operational Flight Data Monitoring
PVT	Psychomotor Vigilance Task
SAFTE	Sleep, Activity, Fatigue, and Task Effectiveness
SP	Samn-Perelli
TEM	Threat and Error Management
WOCL	Window Of Circadian Low

European Commission