Effectiveness of Flight Time Limitation (FTL)

Working Document 2.1 *Identification of Potential Fatigue Hotspots*

Based on survey and roster data

MOVE/C2/2016-360
Classification: Restricted

Table of Contents

Chapter 1: Introduction to the research study

Main objective and scope of the research study

The European Aviation Safety Agency (EASA) was mandated to perform a continuous review of the effectiveness of the rules concerning flight and duty time limitations and rest requirements contained in Annexes II and III of Commission Regulation (EU) No. 965/2012[1].

The review commenced in 2017 with the commission of a research study.

The research study was broken down into smaller phases; each focused on specific flight duty periods (FDPs). The first and current research phase studied the following two FDPs:

- FDP1: Duties of more than 10 hours at the less favourable time of day.
  This focuses on operations that encroach (fully or partially) any portion of the period between 02:00h and 04:59h; and
- FDP2: Disruptive schedules.
  This focuses on consecutive early duty starts, late duty finishes, night duties, and combinations thereof.

Scope of the current working document

This Working Document 2.1 (Identification of Potential Fatigue Hotspots) reports the results of the work performed to identify potential fatigue hotspots in the target population, based on an online survey across Europe and the analysis of historical pilot and cabin crew roster data using two bio-mathematical models.

[1] 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.

Chapter 2: Data analysis methods

Approach in identifying fatigue hotspots

This section provides a rationale specifying the criteria for the identification of the fatigue hotspots. It explains the approach that we defined for the data analysis.

We used an online survey and bio-mathematical modelling to identify fatigue hotspots. The pilot and cabin crew roster data were gathered from the participating airlines. These rosters were analysed using two bio-mathematical models.

Survey

We developed and used an online survey (see Appendix 1: Online survey) to identify potential work patterns that were associated with fatigue. This survey was also used to provide for a subjective ranking of the FDPs of interest as described in D1 (Definition of the Baseline) and the D1 Addendum.

Besides ranking the FDPs of interest, the survey was used to collect aircrew insights about fatigue hotspots. The respondents selected, from a pre-defined list of 'fatigue items', the items that they deemed to be most relevant for causing the fatigue hotspot.

The respondents could also describe in their own words (i.e., answering open questions) how the rosters affected their fatigue, when they felt most fatigued during the duty, and which conditions were worsening their fatigue. These questions were only asked to those respondents who had indicated to have experience with at least one of two FDPs of interest; namely:

- FDP1: Night duties of more than 10 hours; and
- FDP2: Disruptive schedules.

The frequency of indicated fatigue items was calculated and visualised in bar and pie charts (in Chapter 4: Roster results).

Roster

Data on planned and worked rosters of pilots and cabin crew members, spanning approximately one year, were collected from the airlines participating in the data collection. The data analysis focussed mainly on the worked rosters. The planned rosters were used only for comparison to the worked rosters.

Airline rosters were analysed using two bio-mathematical models in order to predict the potential level of fatigue; Boeing Alertness Model (BAM[2]) and Sleep, Activity, Fatigue, and Task Effectiveness, Fatigue Avoidance Scheduling Tool (SAFTE-FAST[3]). The parameter settings in both models were aligned as far as possible (see Appendix 2: Model parameter settings).

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 Karolinska Sleepiness

[2] Jeppesen – as member of the project team – provided access to BAM.
[3] The project team was provided free access to the SAFTE model.

Scale[4] (KSS) 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. For each flight and each crew member BAM was configured to assign a single alertness prediction at top of descent (TOD). TOD was defined at half an hour before wheels on ground.

The output of the SAFTE model provides a percentage of performance effectiveness (Effect) from 0 (low effectiveness) to 100 (high effectiveness). There is an inverse relation between the SAFTE effectiveness scale and the KSS and Samn-Perelli (SP) scale[5]: a SAFTE value of 20 corresponds with a KSS of 9 (very sleepy, great effort to keep awake, fighting sleep); and a SAFTE value of 100 is KSS 1 (extremely alert); and a SAFTE value of 20 corresponds with a SP of 7 (completely exhausted, unable to function effectively); and a SAFTE value of 100 is SP 1 (fully alert, wide awake). For each flight and each crew member the SAFTE model was configured to assign an Effect prediction at TOD.

We performed the analysis using the dependent variables CAS and Effect estimated for TOD during the final leg of the FDP. The data analysis plan consisted of the following steps.

Step 1: Check for high predicted fatigue scores

The goal of this step was to identify whether or not high predicted fatigue scores occurred in FDP1 (Night duties of more than 10 hours) and FDP2 (Disruptive schedules).

The following values defined a high level of predicted fatigue: scores equal to or below 2,500 on the BAM CAS and scores equal to or below 77 on the SAFTE Effect scale; i.e., equivalent to scores 7 or higher on the KSS and 6 or higher on the SP scale. A high level of predicted fatigue was also defined by an Effect score of below 88.5 for a minimum duration of 90 minutes (referred to as TimeLowEffect); i.e., equivalent to a KSS score of 5 or higher, and an SP score of 4 or higher, both for a minimum duration of 90 minutes.

For a detailed description of the dependent (and independent) variables see Appendix 3: Variables list.

<u>FDP Baseline set</u>

This step determines the probabilities of the occurrence of high predicted fatigue scores for an FDP Baseline set. The FDP Baseline set consists of all FDP data points available in the dataset. The baseline probabilities are utilized in the secondary-objective analysis for FDP1 and FDP2 to determine ratios of the occurrence of high predicted fatigue scores during the two FDPs of interest compared to the FDP Baseline set.

[4] 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.

[5] SP is a 7-point scale: 1. Fully alert, wide awake, 2. Very lively, but not at a peak, 3. Okay, somewhat fresh, 4. A little tired, less than fresh, 5. Moderately tired, let down, 6. Extremely tired, very difficult to concentrate, 7. Completely exhausted, unable to function effectively.

Primary objective FDP1 (Night duties of more than 10 hours)

The objective was to assess the prevalence of high predicted fatigue scores during duties of more than 10 hours at the less favourable time of day (between 02:00h and 04:59h). To this end, the following two operational hypotheses are formulated:

H_0 = High predicted fatigue scores do not occur in flight duties longer than 10 hours that take place between 02:00h and 04:59h.
H_1 = High predicted fatigue scores occur in flight duties longer than 10 hours that take place between 02:00h and 04:59h.

The assessment consisted of estimating the probability that the CAS scored 2,500 or lower, or the Effect scored 77 or lower; and the TimeLowEffect scored below 88.5 for a minimum duration of 90 minutes. A point estimate as well as a confidence interval for CAS TOD, Effect TOD and TimeLowEffect occurrence probabilities were determined[6].

Secondary objective FDP1 (Night duties of more than 10 hours)

The objective was to assess whether high predicted fatigue scores occur during duties of more than 10 hours at the less favourable time of day (between 02:00h and 05:00h) more frequently than in the FDP Baseline set. To this end, the following two operational and statistical hypotheses are formulated:

H_0 = High predicted fatigue scores occur equally frequently in night duties longer than 10h that take place between 02:00h and 04:59h as in the FDP Baseline set.
H_1 = High predicted fatigue scores occur more frequently in night duties longer than 10h that take place between 02:00h and 04:59h than in the FDP Baseline set.

The assessment consisted of calculating the ratio for FDP1 compared to the FDP Baseline set. The ratio was defined as the ratio of the occurrence probabilities of high predicted fatigue scores for the two datasets.

Primary objective FDP2 (Disruptive schedules)

The objective was to assess the prevalence of high predicted fatigue scores during consecutive disruptive FDPs. To this end, the following two operational hypotheses are formulated:

H_0 = High predicted fatigue scores do not occur in consecutive disruptive FDPs, irrespective of number of FDPs and type: early start, late finish, night, or mix.
H_1 = High predicted fatigue scores occur in consecutive disruptive FDPs irrespective of number of FDPs and type: early start, late finish, night, or mix.

Secondary objective FDP2 (Disruptive schedules)

The objective was to assess whether or not high predicted fatigue scores during consecutive disruptive FDPs occur more frequently than in the FDP Baseline set. To this end, the following two operational and statistical hypotheses are formulated:

H_0 = High predicted fatigue scores occur equally frequently in consecutive disruptive FDPs, irrespective of number of FDPs and type: early start, late finish, night, or mix, as in the FDP Baseline set.
H_1 = High predicted fatigue scores occur more frequently in consecutive disruptive FDPs irrespective of number of FDPs and type: early start, late finish, night, or mix, than in the FDP Baseline set.

[6] Concerning the point estimate, two cases were distinguished, namely a point estimate equal to zero or a point estimate larger than zero. Zero observed occurrences lead to a formal point estimate of zero but do not imply that the true occurrence rate equals zero. The true occurrence rate (with the specified level of confidence) may be as large as the upper limit of the estimated confidence interval. Non-zero observed occurrences result in a point estimate larger than zero. The corresponding confidence interval may or may not include the value of zero. When the interval does not include zero, the occurrence probability of high levels of fatigue is significant.

Step 2: Find clusters of variables

The goal of this step was to develop multiple regression models that can be used to determine the FDP characteristics under which high levels of fatigue were predicted by the models.

The FDP1 assessment was similar to the FDP2 assessment. The FDP2 assessment focused solely on the different types of *consecutive* duties; i.e., consecutive early starts, consecutive late finishes, consecutive nights, and combinations (mix) thereof.

The variables that may contribute to the prediction of high levels of predicted fatigue (i.e., CAS ≤ 2,500 and Effect ≤ 77) were evaluated in a multiple logistic regression analysis. Logistic regression predicts the occurrence of a binary dependent variable: an event that either takes place (i.e., a high level of fatigue) or does not take place (i.e., a low level of fatigue). Logistic regression models the relationship between a binary dependent variable and one or more nominal, ordinal, interval or ratio-level independent variables. Multinomial logistic regression extends logistic regression to the case where a dependent variable can take more than two (discrete) values[7].

Variables that may contribute to fatigue can be found in Appendix 3: Variables list. These were defined based upon the following sources:

- The online survey findings;
- The parameters in the bio-mathematical models that were used for the analyses of roster data;
- Scientific literature review[8]; and
- Ideas and suggestions from scientific committee and consortium members.

Step 3: Compare planned and worked rosters

The planned roster data was compared with the worked roster data to study the stability of the roster planning and the impact of roster changes on level of predicted fatigue. In order to be able to compare the planned and worked rosters, we paired (or matched) the datasets using flight numbers, departure locations, arrival locations, and departure dates.

Note that the datasets were completely de-identified before delivery. This limited the comparison between planned and worked rosters as we could not look into the individual crew effects (especially on predicted sleep and CAS and Effect scores) of changes in the roster; i.e., there was no way of knowing if the crew that was scheduled for a planned flight actually performed this flight or that the flight was moved over to another crew.

Note that we could not determine when the planned rosters were published in relation to the actual performance of the flights. We requested the participating airlines to provide their latest publication of planned rosters. However, the datasets received did not allow us to look into the question whether or not crews, in case of a late change in the roster, get enough time to prepare for their new schedule, especially in terms of sleep before the flight duty.

[7] For more information on logistic regression analyses check Hosmer, D. W., Lemeshow, S., and Sturdivant, R. X. (2013). *Applied logistic regression*, 3rd edition, Wiley.
[8] As presented in D1 (Definition of the Baseline).

Chapter 3: Survey results

Demographic data

A high-level overview of the demographic data is provided below. A more thorough description of the population of respondents to the survey is provided in D1 Addendum.

The total number of aircrew respondents was 15,680 (28.4% female); i.e., 10.6% of the entire aircrew population base in Europe[9]. Of these respondents 58.2% were pilots (4.5% female) and 41.8% cabin crew members (61.5% female). The mean age of all crew respondents was 41 years and 8 months old (range 17 - 75). The mean age for pilots was 42 years and 4 months and for cabin crew 40 years and 10 months.

Of the aircrew respondents 27.5% indicated to work for a point-to-point operator; 61% worked for a network operator; 3.3% for a cargo operator; and 8.2% for another type of airline.

Fatigue items considered most relevant for causing fatigue

The respondents were asked the following question:

- *Think about the last time you were in active duty (not on a positioning flight) and experienced fatigue you believe was caused by the schedule. Then think about your schedule the days before feeling so fatigued. Please indicate from the list below one or more items that you deem relevant for causing the situation.*

Table 1 Fatigue items

Fatigue item	***Abbreviation***[10]
Starting early	EarlyStart
Finishing late	LateFinish
Long working days	LongDay
Not sleeping at home for several days in a row	Hotel
Outward westward flight across >6 time zones	OutWest
Outward eastward flight across >6 time zones	OutEast
Return flight after a westward flight across >6 time zone	InWest
Return flight after an eastward flight across >6 time zones	InEast
Flying a great number of sectors	Sector
Unfavourable times for resting (in period when you are not sleepy)	BadRestTime
Short recovery time between duties	ShortReco
Insufficient rest time during flight	InsRestTime
Insufficient quality of on-board rest facilities	BadQual
Flying during hours when I would normally sleep	WorkSleep

Figure 1 shows how frequently pilots and cabin crew indicated the specific fatigue items.

[9] As estimated in D2.2 (Definition of the Data Collection Process).
[10] In Table 1, next to each fatigue item, the abbreviation for this item was given. These abbreviations were used in the graphs below.

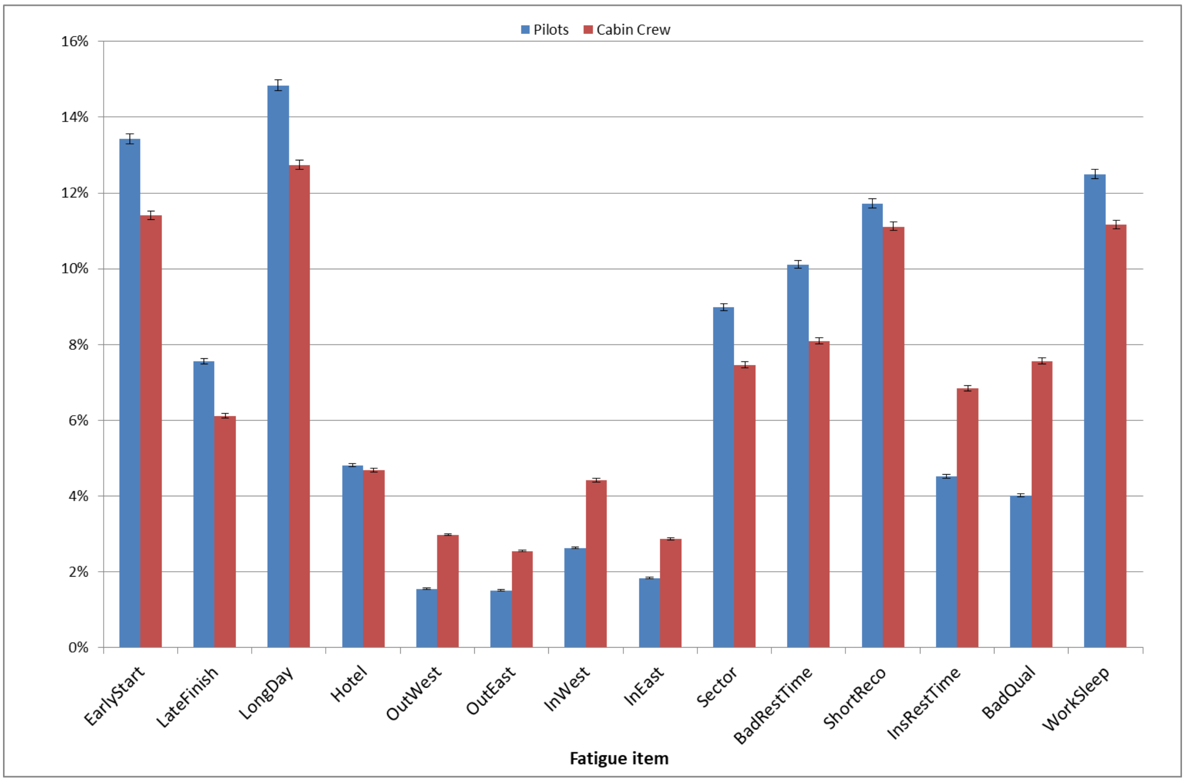

Figure 1 Frequency of indicated fatigue items by pilots and cabin crew. The error bars indicate the 95% confidence interval

Figure 1 shows a similar pattern for pilots and cabin crew. The respondents indicated early starts, long days, short recovery time between duties, and flying at times when you normally sleep most frequent. They also often referred to flying a high number of sectors and short recovery time between duties.

Figure 2 shows how the respondents from different regions of Europe rated the fatigue items. Four geographical regions were defined covering Europe: East, West, North, and South (see Table 2). The figure shows small differences in region for early starts and number of sectors; respondents from the Northern and Eastern part of Europe indicated these items most frequently. In North Europe respondents referred to long days as a fatigue item more than in the rest of Europe.

Table 2 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		

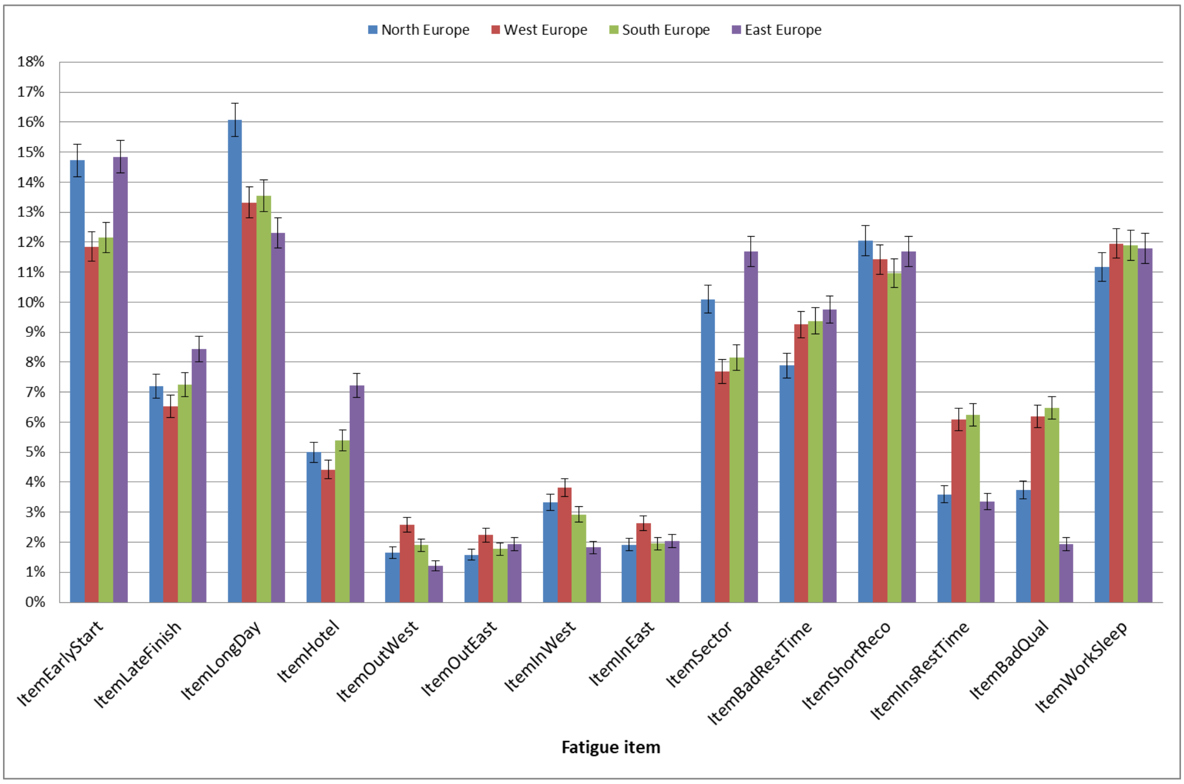

Figure 2 Selected fatigue items per area of Europe. The error bars indicate the 95% confidence interval

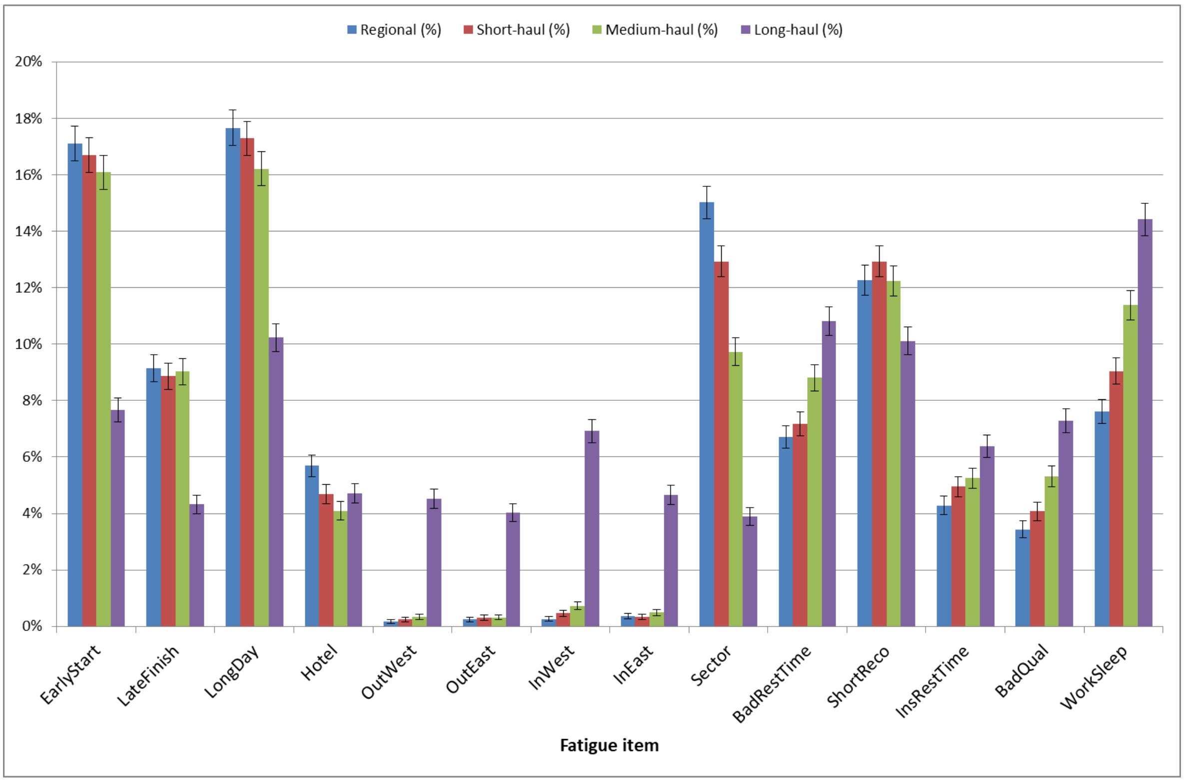

Figure 3 Selected fatigue items per haul type. The error bars indicate the 95% confidence interval

Figure 3 shows how the respondents flying different haul types rated the fatigue items. The figure shows quite some differences between the fatigue items indicated by the respondents per haul type. Respondents indicated early starts, long days, and high number of sectors as fatigue items most frequently for regional, short- and medium-haul operations; long-haul operations was referred to relatively less here.

Results from open questions

The survey contained a number of open questions to obtain information on how fatigue was affected:

- Please describe how the roster preceding FDP1 or FDP2 affects your fatigue.
- Concerning FDP1 or FDP2, when do you feel most fatigued and what makes this type of duty so fatiguing?
- Which conditions may worsen fatigue in FDP1 or FDP2?

Only respondents who indicated to have experience with either FDP1 or FDP2 could answer the open questions. Therefore, the number of respondents who completed this part of the survey is a subset of the total number of respondents.

Due to the large response to the open questions, we decided to use a randomly selected 10% of the responses in the analysis. Responses to questions on the two FDPs were clustered according to similarity. Respondents could provide several reasons/arguments in one answer. The number of responses was therefore not necessarily the same as the number of respondents. Note that the category 'no answer provided' was not included, and the category 'others' comprises of answers that were given but did not fit within one of the main categories.

Different respondents used different words to make their statements. For analysis purposes these answers were aggregated into several categories to get a better overview of the types of factors considered most relevant by the respondents.

The aim of this analysis was to identify which factors contribute to being fatigued, in addition to the fatigue items that were already indicated by the respondents in a previous question. The aggregated answers that were given were visualised in pie charts.

Results for FDP1 (Night duties of more than 10 hours)

Figure 4 shows features of a roster preceding a night duty of more than ten hours at the less favourable time of day that contributed to feeling fatigued. The results indicated three main fatigue items for FDP1; i.e. insufficient recovery time in between duties, disruptive schedule (incl. early starts), and flying or being awake during sleeping hours (i.e. in the window of circadian low, WOCL: 02:00h - 05:59h). Thirteen percent of the respondents indicated starting early as a fatigue item, which is remarkable for a night duty.

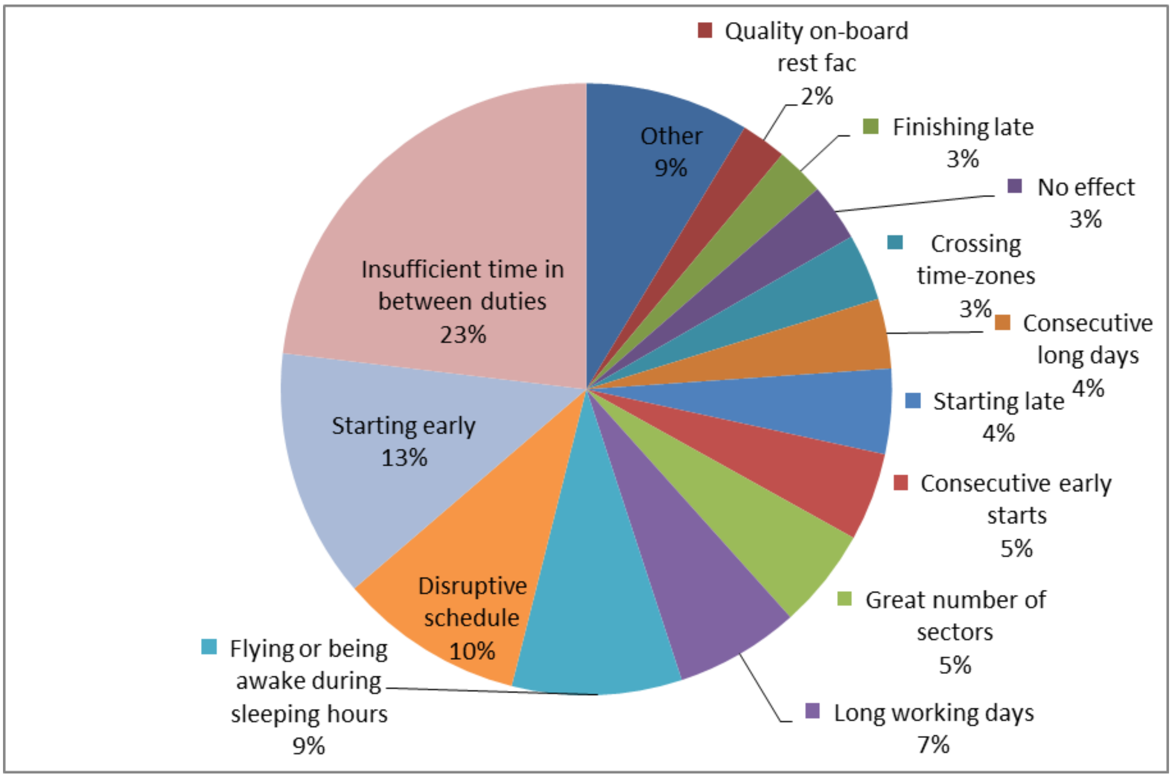

Figure 4 Aggregated responses to: "*Please describe how the roster preceding FDP1 affects your fatigue*". This chart is based upon 623 answers

Figure 5 indicates when respondents felt most fatigued and what the main cause of that was. From this figure it showed that respondents felt most fatigued during the WOCL. A few hours after duty start and (before) the end of duty were also mentioned frequently. Further, respondents indicated being fatigued after a number of consecutive flights.

Figure 6 indicates which conditions worsened fatigue. The most often mentioned item here was the item delays. Respondents reported weather as a factor that worsened fatigue. The item poor rest on the plane referred to the quality of the rest facilities as well as the poor quality of sleep. Both high workload as well as low workload were indicated as fatigue items. The item crew and passengers referred to persons asking difficult questions or otherwise hindering/irritating the crew.

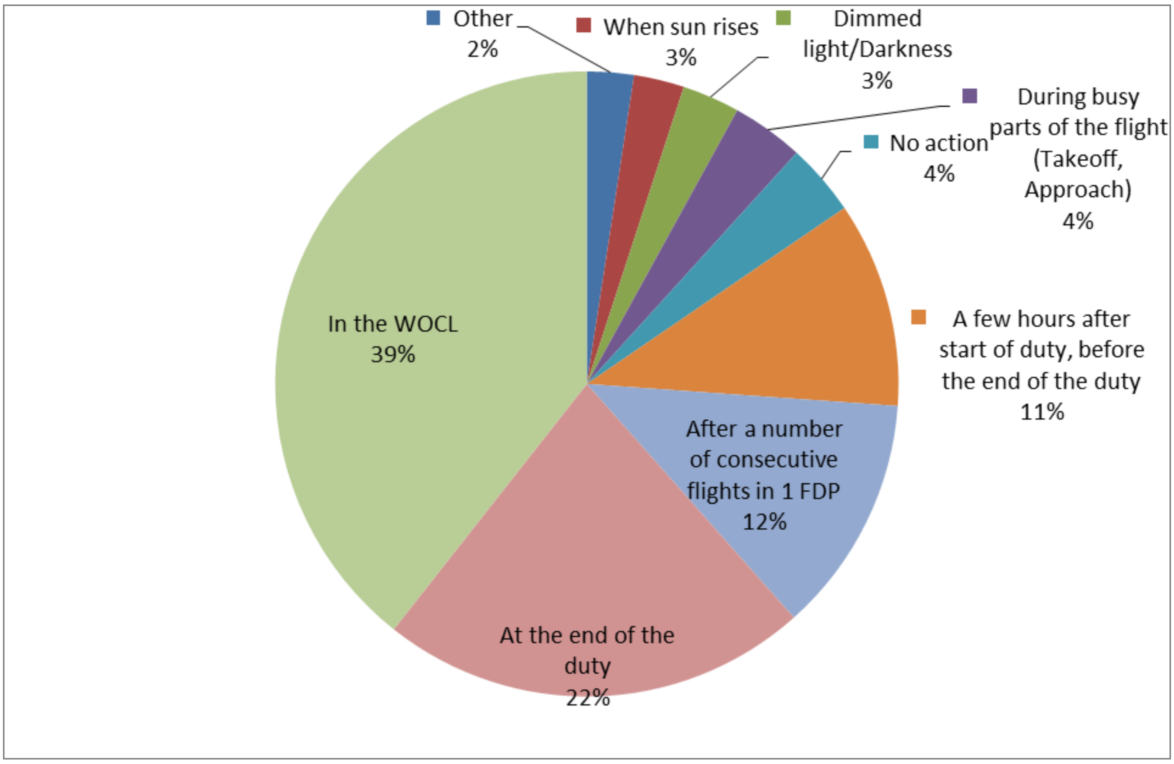

Figure 5 The answers given to: "*Concerning FDP1, when do you feel most fatigued and what makes this type of duty so fatiguing?*". This chart is based upon 536 answers

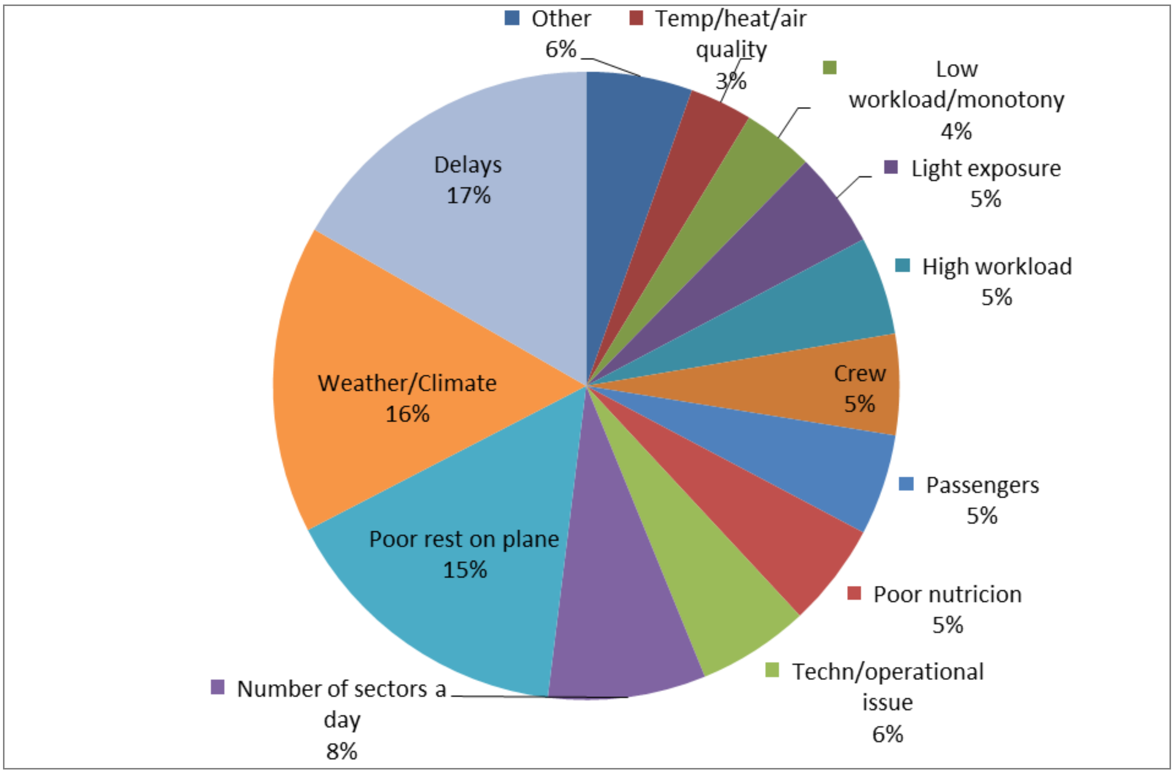

Figure 6 The answers given to: *"Which conditions may worsen fatigue in FDP1?"*. This chart is based upon 712 answers

Results for FDP2 (Disruptive schedules)

Figure 7 indicates what aspects of a roster contributed to fatigue for FDP2 according to the respondents. The results indicated insufficient time in between duties as the main fatigue item. Other frequently mentioned items were roster changes, starting early, unfavourable times for resting, long working days, and flying or being awake during sleeping hours.

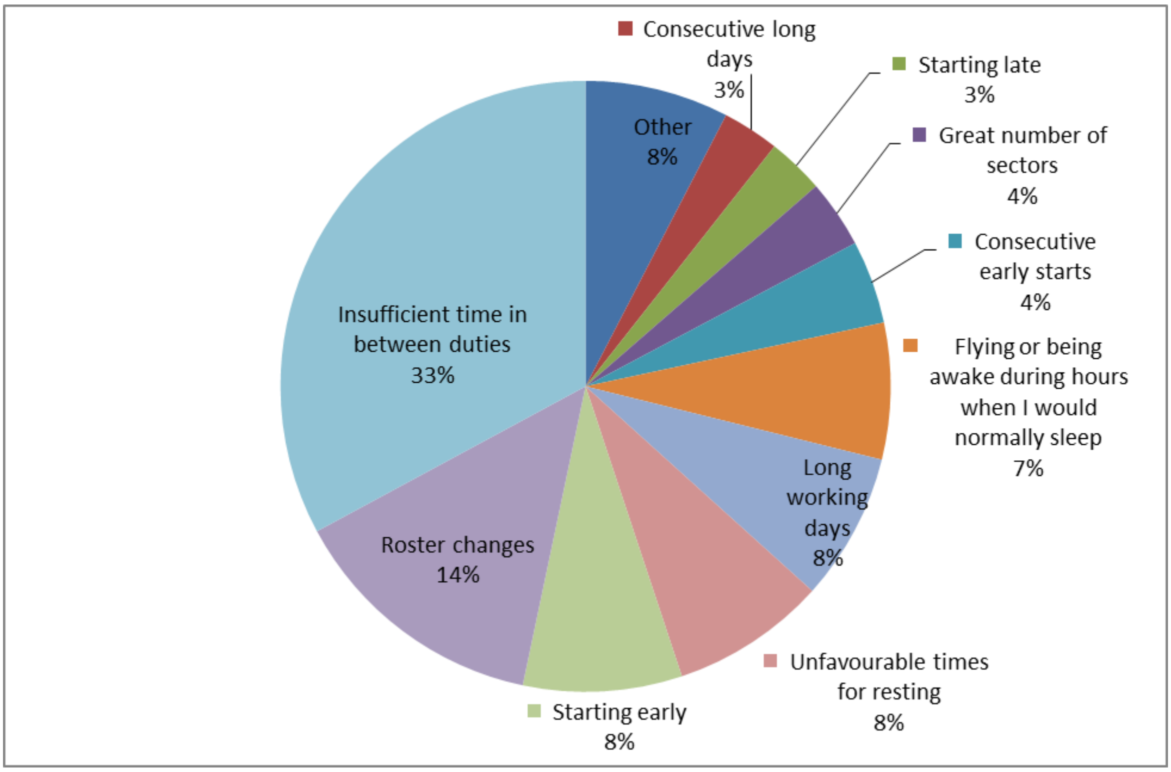

Figure 7 Aggregated responses to: "*Please describe how the roster preceding FDP2 affects your fatigue*". This chart is based upon 499 answers

Figure 8 indicates when respondents felt most fatigued and what the main cause of that was. From this figure it was clear that respondents felt most fatigued during the WOCL and at the end of the duty.

Figure 9 indicates which conditions worsened fatigue. The most often mentioned items that worsened fatigue were delays and weather. The item poor rest on the plane referred to the quality of the rest facilities, as well as a poor quality of sleep. The item crew and passengers referred to persons asking difficult questions or otherwise hindering/irritating the respondent.

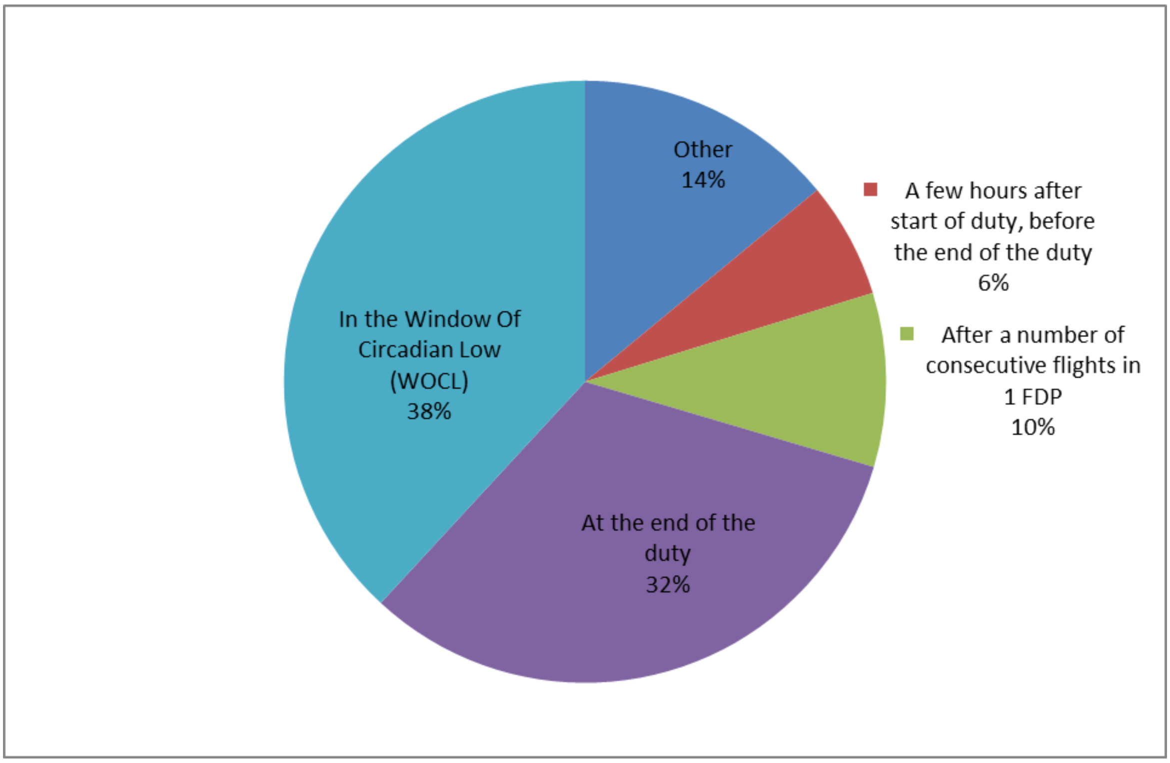

Figure 8 The answers given to: "*Concerning FDP2, when do you feel most fatigued and what makes this type of duty so fatiguing?*". This chart is based upon 257 answers

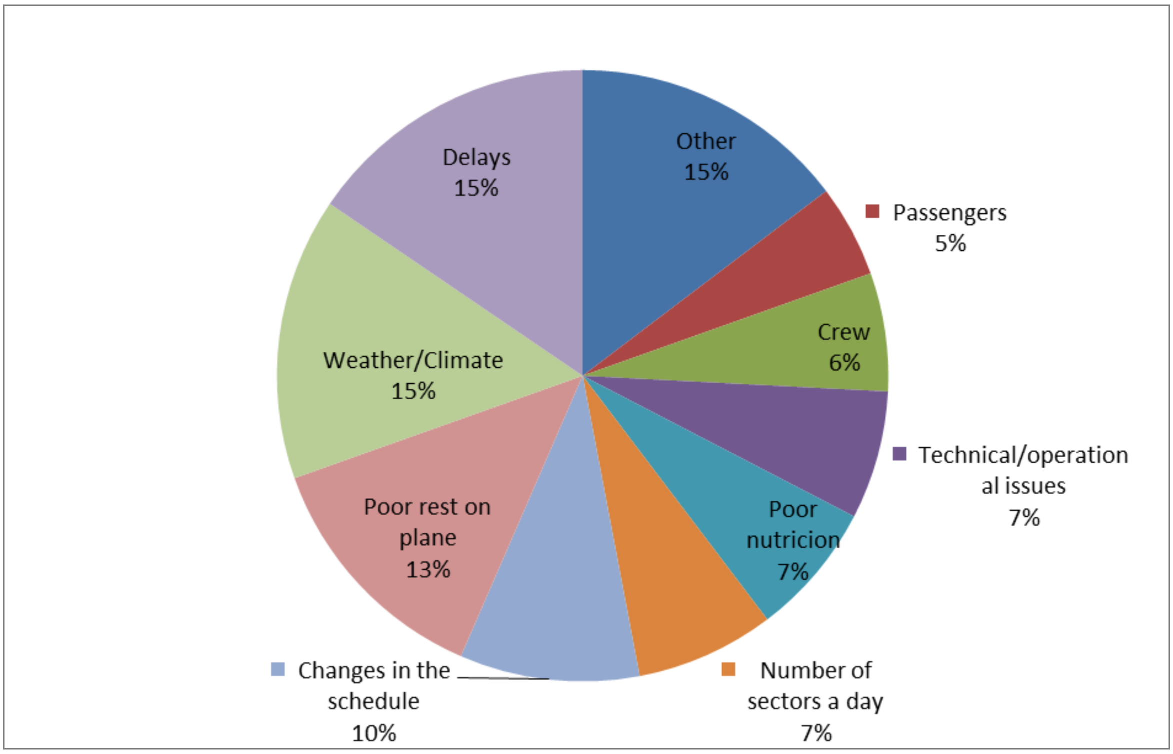

Figure 9 The answers given to: "*Which conditions may worsen fatigue in FDP2?*". This table is based upon 368 answers

Chapter 4: Roster results

Description of worked roster data

Data of about one year on worked rosters of pilots and cabin crew were collected from the airlines participating in the data collection. In total rosters of six airlines were used for the data analysis. Not all rosters that we received from the airlines could be run in the models due to incompleteness, format issues, and time constraints[11]. The six rosters were merged into one roster database in which the FDPs of interest were marked.

Four geographical European regions were defined in D1 Addendum: East, West, North, and South. Table 3 shows how the six airlines divided across these regions.

Table 3 No. of airlines participating across Europe

	East	*West*	*North*	*South*
No. of airlines	2	2	0	2

Five types of operation were defined in D1 Addendum. Table 4 provides an overview of the types of operation of the participating airlines. The total number of airlines exceeded six as most of the airlines operate more than one type of operation.

Table 4 Types of operation of the participating airlines

	Long-haul	*Medium-haul*	*Short-haul*	*Regional*	*Cargo*
No. of airlines	4	4	5	3	1

Table 5 provides information on the number of FDPs per haul type in the FDP Baseline set. The FDP Baseline set consists of all FDP data points available in the dataset; i.e., the FDP data points resulting from the worked rosters.

Table 5 No. of FDPs per haul type

Haul type	*Frequency*	*Percent*
Long-haul (more than 5 h and crossing 3 time zones)	29,471	11.1%
Medium-haul (more than 2 h)	67,251	25.4%
Short-haul (between 1 and 2 h)	142,971	54.0%
Regional (less than 1 h)	12,672	4.8%
Cargo	12,381	4.7%
Total	264,746	100.0%

A total of 264,746 FDPs (i.e., the FDP Baseline set after corrections for non-acclimatisation and crew augmentation) were gathered.

[11] Sixteen airlines offered their rosters to the project: Air Baltic, Air Europa, Alitalia, ASL Airlines Belgium, Cargolux, Czech Airlines, Flybe, Iberia, KLM, Lufthansa, Lufthansa Cargo, Norwegian Air Shuttle, TAROM, Vueling, SunExpress Deutschland, and Adria Airways.

Figure 10 shows the FDPs gathered over time.

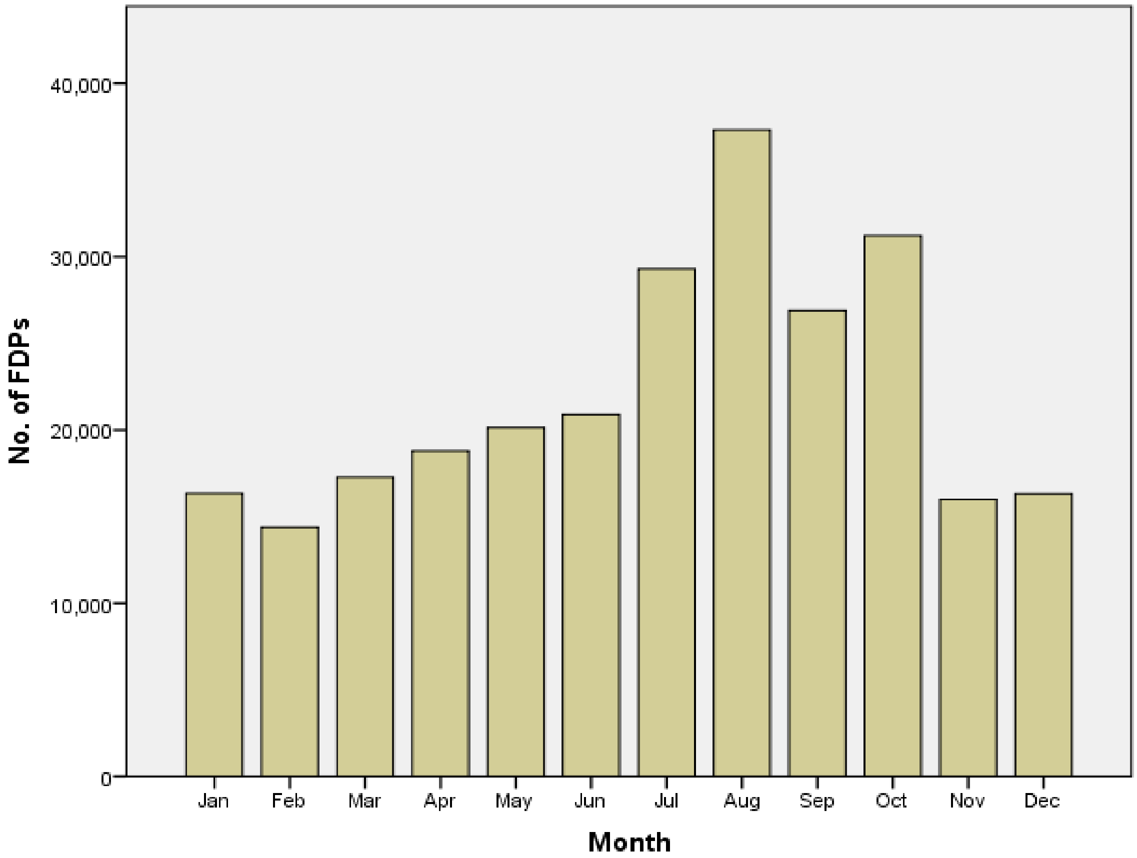

Figure 10 No. of FDPs gathered per month

Table 6 to Table 9 show the numbers of FDPs resulting from the rosters, including the FDPs of interest: FDP1 (Night duties of more than 10 hours) and FDP2 (Disruptive schedules).

Table 6 Sample size FDP1 (Night duties of more than 10 hours), FDP2 (Disruptive schedules), and FDP Baseline

	FDP1 (Night duties > 10h) *No.*	***FDP2 (Disruptive schedules)*** *No.*	***FDP Baseline*** *No.*
Total FDP sample size	12,791	111,951	264,746

Table 7 Sample size FDP2 (Disruptive schedules)

	Early starts *No.*	***Late finishes*** *No.*	***Nights*** *No.*
Total FDP sample size	33,697	30,773	47,481

Table 8 Sample size consecutive FDP2 (Disruptive schedules): Consecutive early starts, late finishes, and nights

	Consecutive early starts *No.*	***Consecutive late finishes*** *No.*	***Consecutive nights*** *No.*
2 in a row	5,494	3,170	1,843
3 in a row	1,119	443	303
4 in a row	241	57	133
5 in a row	34	7	49
6 in a row	7	-	-
Total FDP sample size	6,895	3,677	2,328

Table 9 Sample size FDP2 (Mix)

	Early start - Late finish *No.*	***Late finish - Night*** *No.*	***Night - Early start*** *No.*	***Late finish - Early start*** *No.*	***Night - Late finish*** *No.*	***Early start - Night*** *No.*
Total FDP sample size	1,030	3,332	1,255	1,675	1,236	5

Results for FDP1 (Night duties of more than 10 hours)

Check for high predicted fatigue scores

This section begins with the primary objective for FDP1 (Night duties of more than 10 hours) by presenting point estimates and 95% confidence intervals for the occurrence probabilities of high levels of predicted fatigue. Following that, the same quantities will be presented for the FDP Baseline set. Subsequently, the secondary objective will be addressed by examining the ratio for FDP1 (Night duties of more than 10 hours) compared to the FDP Baseline set.

Results were presented for CAS TOD, Effect TOD and TimeLowEffect. Other than missing values, the outcome measures were not affected by outliers.

<u>Primary objective for FDP1 (Night duties of more than 10 hours)</u>

The objective was to assess the prevalence of high predicted fatigue scores during duties of more than 10 hours at the less favourable time of day (between 02:00h and 04:59h).

The hypotheses were addressed by estimating the probability of the occurrence of high predicted fatigue scores[12]. Table 10 presents occurrence-probability point estimates as well as 95% lower and upper confidence limits for high scores of predicted fatigue resulting from the measures CAS TOD, Effect TOD and TimeLowEffect. The numbers of valid measurements (Nvalid) are included. We concluded that high fatigue scores did occur in the flight duties longer than 10 hours that encroach (part of) the period between 02:00h and 04:59h as postulated under the operational hypothesis H_1.

[12] Note that, with regard to the operational hypothesis H_0, not observing any high predicted fatigue scores in the data does not imply that the underlying occurrence probability is zero. If the observed number of high fatigue scores was zero, then we used the 95% upper confidence limit for the occurrence probability as a conservative point estimate.

Table 10 CAS TOD, Effect TOD and TimeLowEffect occurrence-probability point estimates for FDP1 (Night duties > 10h)

Fatigue measure	N	*p*	*p*L	*p*U
CAS TOD =< 2,500	7,405	0.579	0.570	0.587
Nvalid	12,791			
Effect TOD =< 77	4,891	0.382	0.374	0.391
Nvalid	12,791			
TimeLowEffect => 90 min	8,577	0.671	0.662	0.679
Nvalid	12,791			

Occurrence-probability point estimates (*p*) as well as 95% lower (*p*L) and upper confidence limits (*p*U).

<u>FDP Baseline set</u>

The FDP Baseline set consists of all FDP data points available in the dataset. There were no particular objectives for the FDP Baseline set other than it being used for an assessment of the ratio for FDP1 (as well as for FDP2) compared to the FDP Baseline set. The occurrence probabilities of the different outcome measure values of the FDP Baseline set, together with their 95% upper and lower confidence limits, are presented in a similar manner as for FDP1 above.

Table 11 presents the occurrence-probability point estimates as well as 95% lower and upper confidence limits for CAS TOD, Effect TOD and TimeLowEffect. We concluded that high fatigue scores did occur in the Baseline set of all gathered FDPs.

Table 11 CAS TOD, Effect TOD and TimeLowEffect occurrence-probability point estimates for FDP Baseline set

Fatigue measure	N	*p*	*p*L	*p*U
CAS TOD =< 2,500	40,069	0.151	0.150	0.153
Nvalid	264,744			
Effect TOD =< 77	16,736	0.063	0.062	0.064
Nvalid	264,746			
TimeLowEffect => 90 min	38,324	0.145	0.143	0.146
Nvalid	264,746			

Occurrence-probability point estimates (*p*) as well as 95% lower (*p*L) and upper confidence limits (*p*U).

<u>Secondary objective for FDP1 (Night duties of more than 10 hours)</u>

The objective was to assess whether, or not, high predicted fatigue scores during duties of more than 10 hours at the less favourable time of day (between 02:00h and 04:59h) occur more frequently than in the FDP Baseline set.

The ratio estimates for FDP1 (Night duties of more than 10 hours) and the FDP Baseline set for the CAS, Effect and TimeLowEffect scores were respectively 3.834, 3.800 and 3.837. All estimates were considerably larger than 1.0, showing that FDP1 (Night duties of more than 10 hours) was more prone to high predicted fatigue scores than the FDP Baseline set.

Find clusters of variables

The goal of this step for FDP1 (Night duties of more than 10 hours) was to determine the FDP characteristics under which high levels of predicted fatigue (also referred to as fatigue hotspots) occur by means of multiple logistic regression.

Results were presented for the CAS TOD and Effect TOD measures. All computations were performed in IBM SPSS Statistics 25 unless noted otherwise.

It was expected (and confirmed in a trial run) that, due to the large amount of roster data, most or all of the independent variables would be significant. For practical reasons we decided to constrain the number of independent variables to a top ten subset of variables, where the criterion for selection was based on the magnitude of the correlation coefficient between the dependent variable, CAS or Effect, and the different independent variables, and operational relevance.

Table 12 shows the top ten independent variables with the highest Pearson correlation (in magnitude) within the total set of variables. The Pearson correlation coefficient provides a measure for the linear relationship between two variables. All variables (all in the top ten) showed a significant correlation (based on testing a null hypothesis of zero correlation).

For both outcomes (CAS, Effect), more sleep in the past 24 hours and later FDP end time the second showed the strongest correlations with fatigue at TOD. Also more sleep in the past 48 hours and earlier FDP start time were among the top five independent variables in both analyses (CAS, Effect).

Table 12 Top ten of independent variables for FDP1 (Night duties > 10h) (no adjustments)

***Variables** CAS TOD*	***Rank***	***Correlation coefficient***	***Sample size***
SleepOppD24: Sleep opportunity in darkness in 24h prior TOD	1	0.832[b]	12791
EndH: FDP end time	2	0.808[b]	12791
StartH: FDP start time	3	-0.761[b]	12791
TZ_EW: Time zones crossed from East to West	4	0.667[b]	12791
SleepOppD48: Sleep opportunity in darkness in 48h	5	0.659[b]	12791
TZ_WE: Time zones crossed from West to East	6	-0.576[b]	12791
FDPaltWOCLm: Time in alternative WOCL	7	-0.338[b]	12791
SleepOpp24: Sleep opportunity in 24h prior TOD	8	0.329[b]	12791
SleepOpp48: Sleep opportunity in 48h prior TOD	9	0.283[b]	12791
FDPdur: FDP duration	10	-0.271[b]	12791
***Variables** Effect TOD*	***Rank***	***Correlation coefficient***	***Sample size***
SleepPred24: Sleep prediction in 24h prior TOD	1	0.720[b]	12791
EndH: FDP end time	2	0.707[b]	12791
TimeDayTOD: Time of day at TOD	3	0.703[b]	12791
SleepPred48: Sleep prediction in 48h prior TOD	4	0.702[b]	12791
StartH: FDP start time	5	-0.619[b]	12791
TZ_WE: Time zones crossed from West to East	6	-0.565[b]	12791
TZ_EW: Time zones crossed from East to West	7	0.483[b]	12791
FDPaltWOCLm: Time in alternative WOCL	8	-0.390[b]	12791
FlightDur: Flight duration	9	-0.382[b]	12791
TimeBreak: Time between sectors	10	0.316[b]	12791

[a] Denotes significance of a correlation coefficient at the 5% level.
[b] Denotes significance at the 1% level.

Multiple logistic regression

The following multiple logistic regression model was adopted as a starting point for the regression of the dependent variable Y (CAS and Effect at TOD) on the top ten independent variables X:

$$log\ odds(X) = log\frac{\pi(X)}{1-\pi(X)} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots, \beta_{10} X_{10} \qquad (1)$$

The SPSS enter method was run. This resulted in the inclusion of nine independent variables for both the CAS and Effect variable, and a constant term. Table 13 summarizes the results of this step.

Table 13 Results of multiple regression method for FDP1 (Night duties > 10h)

***Variable** CAS TOD*	***Symbol***	***b***	***OR***	***95% CI for OR***	
				Lower	*Upper*
Constant	X_0	-19.811	.000		
SleepOppD24: Sleep opportunity in darkness in 24h prior TOD	X_1	.000	1.000	1.000	1.000
EndH: FDP end time	X_2	.469	1.598	1.550	1.647
StartH: FDP start time	X_3	-.019	.981	.962	1.000
SleepOppD48: Sleep opportunity in darkness in 48h	X_5	.000	1.000	1.000	1.000
TZ_WE: Time zones crossed from West to East	X_6	-.353	.702	.681	.724
FDPaltWOCLm: Time in alternative WOCL	X_7	-.010	.990	.989	.992
SleepOpp24: Sleep opportunity in 24h prior TOD	X_8	.000	1.000	1.000	1.000
SleepOpp48: Sleep opportunity in 48h prior TOD	X_9	.000	1.000	1.000	1.000
FDPdur: FDP duration	X_{10}	.011	1.011	1.009	1.013
***Variable** Effect TOD*	***Symbol***	***b***	***OR***	***95% CI for OR***	
				Lower	*Upper*
Constant	X_0	4.190	65.992		
SleepPred24: Sleep prediction in 24h prior TOD	X_1	-.015	.985	.983	.986
EndH: FDP end time	X_2	.490	1.632	1.472	1.809
SleepPred48: Sleep prediction in 48h prior TOD	X_4	.008	1.008	1.007	1.009
StartH: FDP start time	X_5	-.037	.964	.945	.982
TZ_WE: Time zones crossed from West to East	X_6	.138	1.148	1.112	1.185
TZ_EW: Time zones crossed from East to West	X_7	.452	1.571	1.504	1.642
FDPaltWOCLm: Time in alternative WOCL	X_8	-.028	.972	.970	.974
FlightDur: Flight duration	X_9	-.010	.990	.989	.991
TimeBreak: Time between sectors	X_{10}	.007	1.007	1.006	1.009

OR Odds Ratio. CI Confidence Interval.

The model parameters can be interpreted as follows: each regression model parameter (i.e., estimate *b*) is the estimated change in the log odds of the response per unit change in the value of the predictor. Positive parameters indicate for every increase in the value of the predictor an increase in the log odds of a high level of fatigue, holding all other variables in the model constant. Conversely, negative parameters mean for each increase in the value of the predictor a decrease in the log odds of a high level of fatigue. The exponential function of the regression model parameter is the odds ratio linked with a one-unit increase in the predictor. It represents the odds that a response (i.e., high level of fatigue) will occur given a specific exposure, compared to the odds of the response occurring in the absence of that exposure. It holds that for an odds ratio value of 1 or, equivalently, a model parameter value of *b* equal to 0, there is no effect of exposure on the odds of response. For OR values > 1 the exposure is associated with higher odds, and for odds ratio values < 1 the exposure is associated with lower odds. The 95% confidence interval is used to estimate the precision of the odds ratio point estimate. It was expected that, due to the large sample size, the confidence interval would be small.

For the CAS measure there is no effect of sleep on the odds of high fatigue. For Effect we see that sleep (in 24h; there is hardly an effect for 48h) decreases the odds of high fatigue. Earlier start and later end time of the duty period show a relatively strong increase in odds in both models. This is confirmed by the significance of the alternative WOCL variable. We see contradicting direction of odds ratios for (more/less) time zones crossed between the models. The FDP/flight duration and time between sectors show hardly an effect on the odds.

Results for FDP2 (Disruptive schedules)

Check for high predicted fatigue scores

This section begins with the primary objective for the duty period FDP2 (Disruptive Schedules) by presenting point estimates and 95% confidence intervals for the occurrence probabilities of high levels of predicted fatigue. Following that, the secondary objective will be addressed by examining the ratio for FDP2 (Disruptive Schedules) compared to the FDP Baseline set.

Results were presented for CAS TOD, Effect TOD and TimeLowEffect. Other than missing values, the outcome measures were not affected by outliers. For FDP2 (Disruptive Schedules), the following subsets were defined in addition to the full FDP2 set:

- FDP2 (Early starts) (i.e., irrespective of the number of repetitions);
- FDP2 (Late finishes);
- FDP2 (Nights);
- Consecutive (i.e., at least two in a row) FDP2 (Early starts);
- Consecutive FDP2 (Late finishes);
- Consecutive FDP2 (Nights); and
- FDP2 (Mix).

These subsets of disruptive FDPs are presented after the full FDP2 set. The FDP2 assessments were similar to the assessment for FDP1 (Night duties of more than 10 hours).

<u>Primary objective for the full FDP2 (Disruptive schedules) set</u>
The objective was to assess the prevalence of high levels of predicted fatigue during disruptive schedule duties, irrespective of the number of repetitions and of type: early start, late finish, or night.

Table 14 presents occurrence-probability point estimates as well as 95% lower and upper confidence limits for CAS, Effect and TimeLowEffect. As can be seen from these estimates, high fatigue scores did occur in disruptive flight schedules as postulated under the operational hypothesis H_1.

Table 14 CAS TOD, Effect TOD and TimeLowEffect occurrence-probability point estimates for the full FDP2 (Disruptive schedules) set

Fatigue measure	N	*p*	*p*L	*p*U
CAS TOD =< 2,500	35,251	0.315	0.312	0.318
Nvalid	111,951			
Effect TOD =< 77	16,104	0.144	0.142	0.146
Nvalid	111,951			
TimeLowEffect => 90 min	35,059	0.313	0.310	0.316
Nvalid	111,951			

Occurrence-probability point estimates (*p*) as well as 95% lower (*p*L) and upper confidence limits (*p*U).

Secondary objective for the full FDP2 (Disruptive schedules) set

The objective was to assess whether or not high predicted fatigue scores during disruptive flight schedules occur more frequently than in the FDP Baseline set.

The ratio estimates for the full FDP2 set and the FDP Baseline set for the CAS, Effect and TimeLowEffect scores were 2.086, 2.286 and 2.159 respectively. All three estimates were considerably larger than 1.0, showing that the full FDP2 (Disruptive schedules) set was more prone to high predicted fatigue scores than the FDP Baseline set.

Primary objective for FDP2 (Early starts)

The objective was to assess the prevalence of high levels of predicted fatigue during early starts, irrespective of the number of repetitions.

Table 15 presents occurrence-probability point estimates as well as 95% lower and upper confidence limits for CAS, Effect and TimeLowEffect. High predicted fatigue scores did occur in the early starts as postulated under hypothesis H_1. However, for Effect, the very small (but non-zero) estimates based on 1 occurrence in 33,697 observations are shown as 0.000 in three decimal places.

Table 15 CAS TOD, Effect TOD and TimeLowEffect occurrence-probability point estimates for FDP2 (Early starts)

Fatigue measure	N	*p*	*p*L	*p*U
CAS TOD =< 2,500	100	0.003	0.002	0.004
Nvalid	33,697			
Effect TOD =< 77	1	0.000	0.000	0.000
Nvalid	33,697			
TimeLowEffect => 90 min	455	0.014	0.012	0.015
Nvalid	33,697			

Occurrence-probability point estimates (*p*) as well as 95% lower (*p*L) and upper confidence limits (*p*U).

Secondary objective for FDP2 (Early starts)

The objective was to assess whether or not high predicted fatigue scores during FDP2 (Early starts) occur more frequently than in the FDP Baseline set.

The ratio estimates for the FDP2 (Early starts) and the FDP Baseline set for the CAS, Effect and TimeLowEffect scores were 0.019, 0.000 and 0.097 respectively. These estimates were considerably smaller than 1.0. This showed that high predicted fatigue scores did not occur more frequently in FDP2 (Early starts) than in the FDP Baseline dataset.

<u>Primary objective for FDP2 (Late finishes)</u>
The objective was to assess the prevalence of high levels of fatigue during late finishes, irrespective of the number of repetitions.

Table 16 presents occurrence-probability point estimates as well as 95% lower and upper confidence limits for CAS, Effect and TimeLowEffect. High predicted fatigue scores did occur in the late finishes as postulated under hypothesis H_1.

Table 16 CAS TOD, Effect TOD and TimeLowEffect occurrence-probability point estimates for FDP2 (Late finishes)

Fatigue measure	N	*p*	*p*L	*p*U
CAS TOD =< 2,500	8,051	0.262	0.257	0.267
Nvalid	30,773			
Effect TOD =< 77	1,885	0.061	0.059	0.064
Nvalid	30,773			
TimeLowEffect => 90 min	3,489	0.113	0.110	0.117
Nvalid	30,773			

Occurrence-probability point estimates (*p*) as well as 95% lower (*p*L) and upper confidence limits (*p*U).

<u>Secondary objective for FDP2 (Late finishes)</u>
The objective was to assess whether or not high predicted fatigue scores during FDP2 (Late finishes) occur more frequently than in the FDP Baseline set.

The ratio estimates for the FDP2 (Late finishes) and the FDP Baseline set for the CAS, Effect and TimeLowEffect scores were 1.735, 0.968 and 0.779 respectively. Where the CAS estimate was larger than 1.0, we found smaller (below 1.0) estimates for Effect and TimeLowEffect. This showed that – depending on the fatigue measure – high predicted fatigue scores did or did not occur more frequently in FDP2 (Late finishes) than in the FDP Baseline dataset.

<u>Primary objective for FDP2 (Nights)</u>
The objective was to assess the prevalence of high levels of predicted fatigue during nights, irrespective of the number of repetitions.

Table 17 presents occurrence-probability point estimates as well as 95% lower and upper confidence limits for CAS, Effect and TimeLowEffect. High predicted fatigue scores did occur in the nights as postulated under hypothesis H_1.

Table 17 CAS TOD, Effect TOD and TimeLowEffect occurrence-probability point estimates for FDP2 (Nights)

Fatigue measure	N	*p*	*p*L	*p*U
CAS TOD =< 2,500	27,100	0.571	0.566	0.575
Nvalid	47,481			
Effect TOD =< 77	14,218	0.299	0.295	0.304
Nvalid	47,481			
TimeLowEffect => 90 min	31,115	0.655	0.651	0.660
Nvalid	47,481			

Occurrence-probability point estimates (*p*) as well as 95% lower (*p*L) and upper confidence limits (*p*U).

Secondary objective for FDP2 (Nights)
The objective was to assess whether or not high predicted fatigue scores during FDP2 (Nights) occur more frequently than in the FDP Baseline set.

The ratio estimates for the FDP2 (Nights) and the FDP Baseline set for the CAS, Effect and TimeLowEffect scores were 3.781, 4.746 and 4.517 respectively. All three estimates were larger than 1.0. This showed that high predicted fatigue scores did occur more frequently in FDP2 (Nights) than in the FDP Baseline dataset.

Primary objective for consecutive FDP2 (Early starts)
The objective was to assess the prevalence of high levels of predicted fatigue during consecutive (i.e., at least two in a row) early starts.

Table 18 presents occurrence-probability point estimates as well as 95% lower and upper confidence limits for CAS, Effect and TimeLowEffect. High predicted fatigue scores did occur in the consecutive early starts for CAS and TimeLowEffect as postulated under hypothesis H_1, but not for Effect. In addition, for CAS, the very small estimates based on 1 occurrence in 6,895 observations are shown as 0.000 in three decimal places for the point estimate and the lower 95% confidence limit.

Table 18 CAS TOD, Effect TOD and TimeLowEffect occurrence-probability point estimates for consecutive FDP2 (Early starts)

Fatigue measure	N	*p*	*p*L	*p*U
CAS TOD =< 2,500	1	0.000	0.000	0.001
Nvalid	6,895			
Effect TOD =< 77	0	0.000	0.000	0.001
Nvalid	6,895			
TimeLowEffect => 90 min	40	0.006	0.004	0.008
Nvalid	6,895			

Occurrence-probability point estimates (*p*) as well as 95% lower (*p*L) and upper confidence limits (*p*U).

Secondary objective for consecutive FDP2 (Early starts)
The objective was to assess whether or not high predicted fatigue scores during consecutive FDP2 (Early starts) occur more frequently than in the FDP Baseline set.

The ratio estimates for the consecutive FDP2 (Early starts) and the FDP Baseline set for the CAS, Effect and TimeLowEffect scores were 0.000, 0.000 and 0.041 respectively. All three estimates were smaller than 1.0, showing that high predicted fatigue scores did not occur more frequently in consecutive FDP2 (Early starts) than in FDP Baseline.

Primary objective for consecutive FDP2 (Late finishes)
The objective was to assess the prevalence of high levels of predicted fatigue during consecutive (i.e., at least two in a row) late finishes.

Table 19 presents occurrence-probability point estimates as well as 95% lower and upper confidence limits for CAS, Effect and TimeLowEffect. High predicted fatigue scores did occur in the consecutive late finishes as postulated under hypothesis H_1 for CAS and TimeLowEffect, but not for Effect.

Table 19 CAS TOD, Effect TOD and TimeLowEffect occurrence-probability point estimates for consecutive FDP2 (Late finishes)

Fatigue measure	N	*p*	*p*L	*p*U
CAS TOD =< 2,500	1,161	0.316	0.301	0.331
Nvalid	3,677			
Effect TOD =< 77	0	0.000	0.000	0.001
Nvalid	3,677			
TimeLowEffect => 90 min	4	0.001	0.000	0.003
Nvalid	3,677			

Occurrence-probability point estimates (*p*) as well as 95% lower (*p*L) and upper confidence limits (*p*U).

The figure below presents the percentages (and confidence intervals) of high predicted fatigue scores (CAS only) for the number of consecutive late finishes.

<u>Secondary objective for consecutive FDP2 (Late finishes)</u>

The objective was to assess whether or not high predicted fatigue scores during consecutive FDP2 (Early starts) occur more frequently than in the FDP Baseline set.

The ratio estimates for consecutive FDP2 (Late finishes) and the FDP Baseline set for CAS, Effect and TimeLowEffect were 2.093, 0.000 and 0.007 respectively. The CAS estimate was larger than 1.0, showing that consecutive FDP2 (Late finishes) was more prone to high predicted fatigue scores than the FDP Baseline set. This was not the case for Effect and TimeLowEffect. Here we found estimates considerably smaller than 1.0.

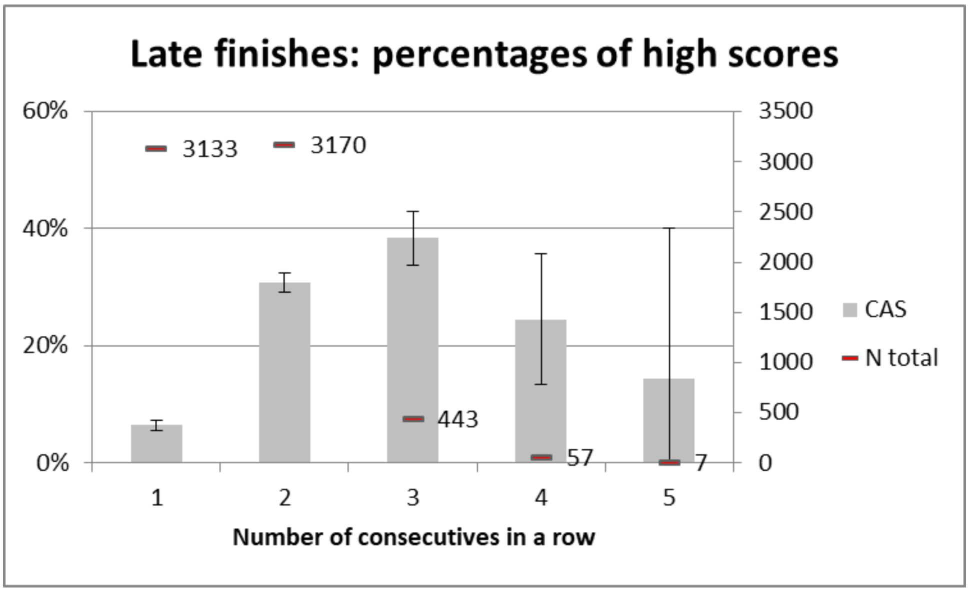

Figure 11 High fatigue scores TOD for consecutive FDP2 (Late finishes)

<u>Primary objective for consecutive FDP2 (Nights)</u>

The objective was to assess the prevalence of high levels of predicted fatigue during consecutive (i.e., at least two in a row) nights.

Table 20 presents occurrence-probability point estimates as well as 95% lower and upper confidence limits for CAS, Effect and TimeLowEffect. High predicted fatigue scores did occur in the consecutive nights as postulated under hypothesis H_1.

Table 20 CAS TOD, Effect TOD and TimeLowEffect occurrence-probability point estimates for consecutive FDP2 (Nights)

Fatigue measure	N	*p*	*p*L	*p*U
CAS TOD =< 2,500	1,935	0.831	0.815	0.846
Nvalid	2,328			
Effect TOD =< 77	943	0.405	0.385	0.425
Nvalid	2,328			
TimeLowEffect => 90 min	1,720	0.739	0.721	0.756
Nvalid	2,328			

Occurrence-probability point estimates (*p*) as well as 95% lower (*p*L) and upper confidence limits (*p*U).

The figure below presents the percentages (and confidence intervals) of high predicted fatigue scores (CAS and Effect) for the number of consecutive nights.

Secondary objective for consecutive FDP2 (Nights)

The objective was to assess whether or not high predicted fatigue scores during consecutive FDP2 (Nights) occur more frequently than in the FDP Baseline set.

The ratio estimates for and the consecutive FDP2 (Nights) and the FDP Baseline set for the CAS, Effect and TimeLowEffect scores were 5.503, 6.429 and 5.097 respectively. All three estimates were larger than 1.0, showing that consecutive FDP2 (Nights) was more prone to high predicted fatigue scores than the FDP Baseline set.

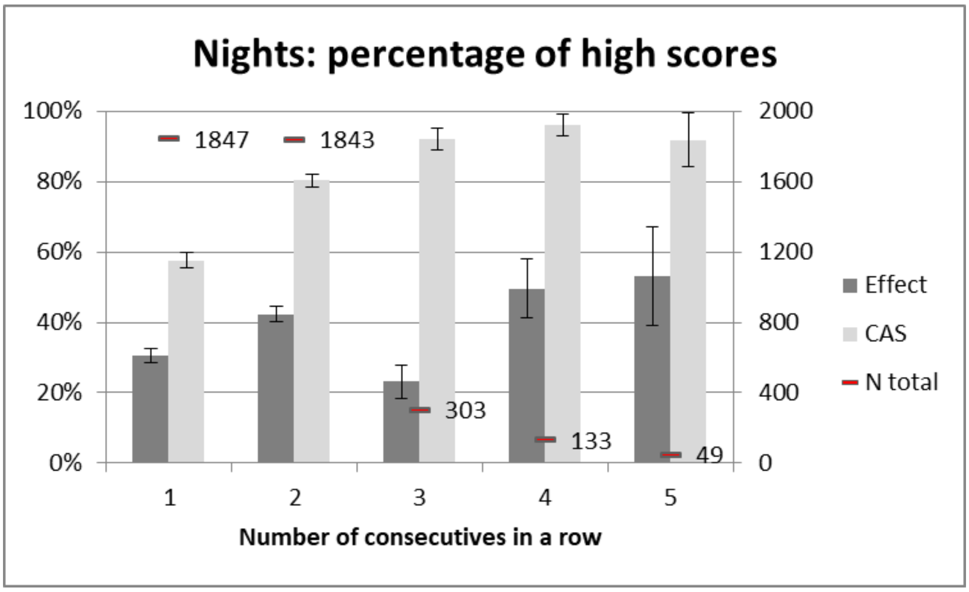

Figure 12 High fatigue scores TOD for consecutive FDP2 (Nights)

Primary objective for FDP2 (Mix)

The objective was to assess the prevalence of high levels of predicted fatigue during a mix of early starts, late finishes and nights.

Table 21 presents occurrence-probability point estimates as well as 95% lower and upper confidence limits for CAS, Effect and TimeLowEffect. High predicted fatigue scores did occur in the consecutive Mix as postulated under the operational hypothesis H_1.

Table 21 CAS TOD, Effect TOD and TimeLowEffect occurrence-probability point estimates for FDP2 (Mix)

Fatigue measure	N	*p*	*p*L	*p*U
CAS TOD =< 2,500	8,401	0.531	0.523	0.539
Nvalid	15,820			
Effect TOD =< 77	7,307	0.462	0.454	0.470
Nvalid	15,820			
TimeLowEffect => 90 min	2,667	0.169	0.163	0.174
Nvalid	15,820			

Occurrence-probability point estimates (*p*) as well as 95% lower (*p*L) and upper confidence limits (*p*U).

The two figures below present the percentages of high predicted fatigue scores (CAS and Effect) for the different transitions into the late finishes and nights respectively. High scores for transitions into early starts hardly occurred (a total of 39 occurrences out of 13,591 for CAS and 3 occurrences out of 13,591 for Effect).

<u>Secondary objective for FDP2 (Mix)</u>
The objective was to assess whether or not high predicted fatigue scores during FDP2 (Mix) occur more frequently than in the FDP Baseline set.

The ratio estimates for FDP2 (Mix) and the FDP Baseline set for the CAS, Effect and TimeLowEffect scores were respectively 3.517, 7.333 and 1.166. All three estimates were larger than 1.0, showing that consecutive FDP2 (Mix) was more prone to high predicted fatigue scores than the FDP Baseline set.

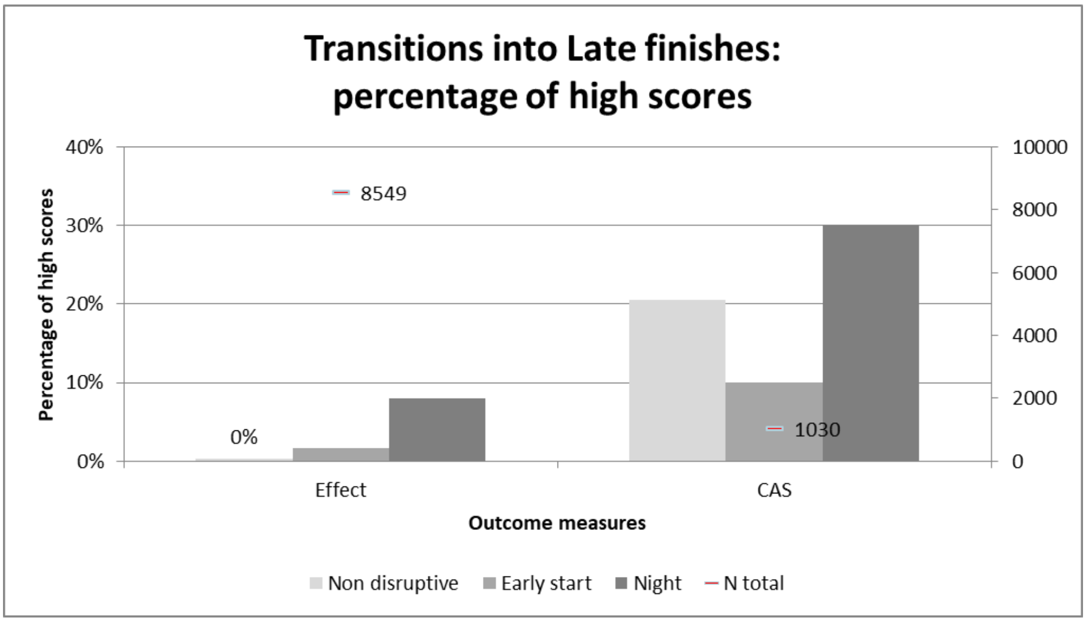

Figure 13 High fatigue scores TOD for transitions into late finishes

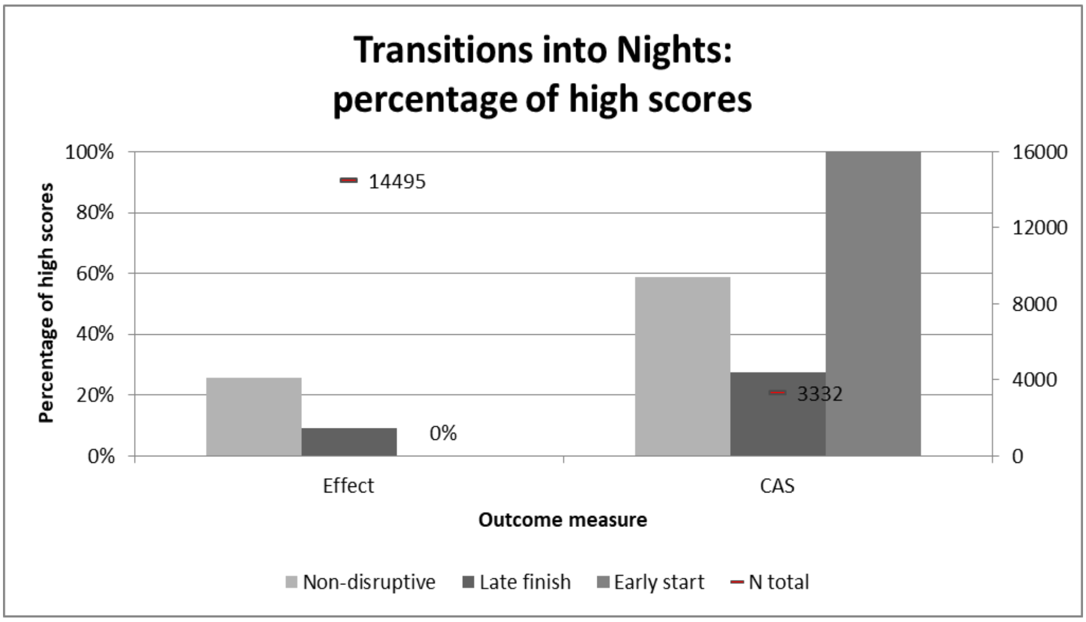

Figure 14 High fatigue scores TOD for transitions into nights

Find clusters of variables

The goal of this step was to determine the FDP characteristics under which high levels of predicted fatigue occur. Results were presented for the CAS TOD and Effect TOD measures.

For this step the focus was on the different types of consecutive FDP2 (Disruptive schedules):

- Consecutive (i.e., at least two in a row) FDP2 (Early starts);
- Consecutive FDP2 (Late finishes);
- Consecutive FDP2 (Nights); and
- FDP2 (Mix).

We decided to constrain the number of independent variables to a top ten subset of variables, where the criterion for selection was based on the magnitude of the correlation coefficient between the dependent variable, CAS or Effect, and the different independent variables, and operational relevance.

Consecutive FDP2 (Early starts)

Table 22 shows the top ten independent variables with the highest Pearson correlation (in magnitude) within the total set of variables. All variables (all in the top ten) showed a significant correlation.

Table 22 Top ten of independent variables for consecutive FDP2 (Early starts) (no adjustments)

***Variables** CAS TOD*	***Rank***	***Correlation coefficient***	***Sample size***
StartH: FDP start time	1	0.677[b]	6895
VeryES: Very early start	2	-0.677[b]	6895
ES: Early start	3	0.677[b]	6895
FDPWOCLm: Time in WOCL	4	-0.670[b]	6895
FDPWOCL%: Percentage in WOCL	5	-0.544[b]	6895
RestPeriod: Rest period	6	0.391[b]	6895
SleepOpp24: Sleep opportunity in 24h prior TOD	7	0.342[b]	6895
SleepOpp48: Sleep opportunity in 48h prior TOD	8	0.328[b]	6895
TZ_EW: Time zones crossed from East to West	9	0.287[b]	6895
TZ_WE: Time zones crossed from West to East	10	0.269[b]	6895
***Variables** Effect TOD*	***Rank***	***Correlation coefficient***	***Sample size***
OutOfPhase: Out of phase	1	0.548[b]	6895
FDPWOCLm: Time in WOCL	2	-0.480[b]	6895
StartH: FDP start time	3	0.477[b]	6895
VeryES: Very early start	4	-0.477[b]	6895
ES: Early start	5	0.477[b]	6895
SleepPred24: Sleep prediction in 24h prior TOD	6	0.442[b]	6895
FDPWOCL%: Percentage in WOCL	7	-0.355[b]	6895
SleepPred48: Sleep prediction in 48h prior TOD	8	0.241[b]	6895
TZ_EW: Time zones crossed from East to West	9	0.207[b]	6895
TZ_WE: Time zones crossed from West to East	10	0.190[b]	6895

[a] Denotes significance of a correlation coefficient at the 5% level.
[b] Denotes significance at the 1% level.

For both outcomes (CAS, Effect), early start time and the time spent inside the WOCL were among the top five independent variables. Note that also the time spent inside the WOCL indicates early FDP start time in this case. Sleep in the past 24/48 hours and the number of time zones crossed were among the lowest five significant independent variables in both analyses (CAS, Effect).

Consecutive FDP2 (Late finishes)

Table 23 shows the top ten independent variables with the highest Pearson correlation (in magnitude) within the total set of variables. All variables showed a significant correlation.

Table 23 Top ten of independent variables for consecutive FDP2 (Late finishes) (no adjustments)

Variables *CAS TOD*	***Rank***	***Correlation coefficient***	***Sample size***
LF: Late finish	1	0.651[b]	3677
VeryLF: Very late finish	2	-0.651[b]	3677
EndH: FDP end time	3	0.625[b]	3677
FlightDur: Flight duration	4	-0.547[b]	3677
FDPdur: FDP duration	5	-0.437[b]	3677
MinPassedTOD: Duty time passed at TOD	6	-0.437[b]	3677
TZ_WE: Time zones crossed from West to East	7	-0.436[b]	3677
SleepOpp48: Sleep opportunity in 48h prior TOD	8	0.374[b]	3677
SleepOppD48: Sleep opportunity in darkness in 48h prior TOD	9	0.289[b]	3677
AwakeTOD: Time awake prior TOD	10	-0.274[b]	3677
Variables *Effect TOD*	***Rank***	***Correlation coefficient***	***Sample size***
LF: Late finish	1	0.724[b]	3677
VeryLF: Very late finish	2	-0.724[b]	3677
TimeDayTOD: Time of day at TOD	3	0.705[b]	3677
EndH: FDP end time	4	0.651[b]	3677
TZ_WE: Time zones crossed from West to East	5	-0.385[b]	3677
FlightDur: Flight duration	6	-0.312[b]	3677
FDPdur: FDP duration	7	-0.195[b]	3677
MinPassedTOD: Duty time passed at TOD	8	-0.195[b]	3677
SleepPred48: Sleep prediction in 48h prior TOD	9	0.136[b]	3677
AwakeTOD: Time awake prior TOD	10	-0.130[b]	3677

[a] Denotes significance of a correlation coefficient at the 5% level.
[b] Denotes significance at the 1% level.

Late FDP finish time was the only independent variable that was among the top five in both analyses (CAS, Effect). FDP/flight duration was among the top five in the CAS analysis and close to that ranking in the Effect analysis, too. Sleep in the past 48 hours and time awake prior to TOD were among the lowest five significant independent variables in both analyses.

Consecutive FDP2 (Nights)

Table 24 shows the top ten independent variables with the highest Pearson correlation (in magnitude) within the total set of variables. All variables (all in the top ten) showed a significant correlation.

Sleep in the past 24 hours and FDP start time were among the top five independent variables in both analyses (CAS, Effect). The time spent inside the WOCL was also close to that ranking. FDP duration and the duty time passed at TOD were among the least five significant independent variables in both analyses.

Table 24 Top ten of independent variables for consecutive FDP2 (Nights) (no adjustments)

***Variables** CAS TOD*	*Rank*	*Correlation coefficient*	*Sample size*
SleepOppD24: Sleep opportunity in darkness in 24h prior TOD	1	0.857[b]	2328
StartH: FDP start time	2	-0.829[b]	2328
SleepOppD48: Sleep opportunity in darkness in 48h prior TOD	3	0.670[b]	2328
SleepOpp24: Sleep opportunity in 24h prior TOD	4	0.620[b]	2328
FDPaltWOCLm: Time in alternative WOCL	5	-0.517[b]	2328
RestPeriod: Rest period	6	0.507[b]	2328
FDPdur: FDP duration	7	-0.504[b]	2328
MinPassedTOD: Duty time passed at TOD	8	-0.504[b]	2328
EndH: FDP end time	9	-0.431[b]	2328
FDPWOCLm: Time in WOCL	10	-0.387[b]	2328
***Variables** Effect TOD*	***Rank***	***Correlation coefficient***	***Sample size***
FlightDur: Flight duration	1	-0.577[b]	2328
StartH: FDP start time	2	-0.449[b]	2328
TZ_WE: Time zones crossed from West to East	3	-0.420[b]	2328
SleepPred24: Sleep prediction in 24h prior TOD	4	0.409[b]	2328
AwakeTOD: Time awake prior TOD	5	-0.350[b]	2328
FDPaltWOCLm: Time in alternative WOCL	6	-0.342[b]	2328
FDPdur: FDP duration	7	-0.338[b]	2328
MinPassedTOD: Duty time passed at TOD	8	-0.338[b]	2328
TZ_EW: Time zones crossed from East to West	9	-0.322[b]	2328
SleepPred48: Sleep prediction in 48h prior TOD	10	0.291[b]	2328

[a] Denotes significance of a correlation coefficient at the 5% level.
[b] Denotes significance at the 1% level.

FDP2 (Mix)

Table 25 shows the top ten independent variables with the highest Pearson correlation (in magnitude) within the total of variables. All variables (all in the top ten) showed a significant correlation.

Only the FDP type called early start (FDP start time between 05:00h - 05:59h or 05:00h - 06:59h depending on country) was among the top five independent variables in both analyses (CAS, Effect). Also, certain duty transitions were among the top five independent variables in both analyses but they were not the same transition types (from night to late finish in the CAS analysis and from late finish to night in the Effect analysis). Early FDP start time defined as 06:00h - 06:59h was among the least five significant independent variables in both analyses.

Table 25 Top ten of independent variables for FDP2 (Mix) (no adjustments)

***Variables** CAS TOD*	***Rank***	***Correlation coefficient***	***Sample size***
SleepOppD24: Sleep opportunity in darkness in 24h prior TOD	1	0.750[b]	8,401
StartH: FDP start time	2	-0.726[b]	8,401
EndH: FDP end time	3	0.704[b]	8,401
FDP2ES: FDP2 early start	4	0.673[b]	8,401
N2LF: Transition from night to late finish	5	-0.576[b]	8,401
FDP2LF: FDP2 late finish	6	-0.570[b]	8,401
RestPeriod: Rest period	7	0.568[b]	8,401
VeryLF: Very late finish	8	-0.487[b]	8,401
SleepOpp24: Sleep opportunity in 24h prior TOD	9	0.486[b]	8,401
ES: Early start	10	0.445[b]	8,401
***Variables** Effect TOD*	***Rank***	***Correlation coefficient***	***Sample size***
FDPaltWOCLm: Time in alternative WOCL	1	-0.713[b]	7,307
FDP2ES: FDP2 early start	2	0.670[b]	7,307
FDPWOCLm: Time in WOCL	3	-0.657[b]	7,307
FDPaltWOCL1h: At least 1h in alternative WOCL	4	-0.589[b]	7,307
LF2N: Transition from late finish to night	5	-0.576[b]	7,307
FDPWOCL%: Percentage in WOCL	6	-0.531[b]	7,307
EndH: FDP end time	7	0.508[b]	7,307
N2ES: Transition from night to early start	8	0.457[b]	7,307
TimeDayTOD: Time of day at TOD	9	0.451[b]	7,307
ES: Early start	10	0.446[b]	7,307

[a] Denotes significance of a correlation coefficient at the 5% level.
[b] Denotes significance at the 1% level.

Multiple logistic regression

The multiple logistic regression model of eq. (1) in the multiple regression section for FDP1 (Night duties of more than 10 hours) was also used for the different subsets of consecutive disruptive FDPs. The SPSS enter method was run for all the subsets.

Consecutive FDP2 (Early starts)
There was only a single valid case; i.e., high fatigue/low CAS score, for modelling of the dependent variable CAS, see Table 18. As a result, model parameter estimation failed.

There were no valid cases; i.e., high fatigue/low Effect scores, for modelling of the dependent variable Effect, cf. Table 18. Statistics could not be computed.

Consecutive FDP2 (Late finishes)
This resulted in the inclusion of nine independent variables for the CAS dependent variable, since one independent variable was excluded by the regression method due to redundancies, and a constant term. Table 26 summarizes the results of this step.

There were no valid cases; i.e., high fatigue/low Effect scores, for modelling of the dependent variable Effect, see Table 19. Statistics could not be computed.

Table 26 Results of multiple regression method for consecutive FDP2 (Late finishes)

Variable *CAS TOD*	***Symbol***	***b***	***OR***	***95% CI for OR***	
				Lower	*Upper*
Constant	X_0	-7.388	.001		
LF: Late finish	X_1	3.234	25.390	17.773	36.273
VeryLF: Very late finish	X_2	.222	1.248	1.222	1.275
FlightDur: Flight duration	X_4	-.009	.991	.988	.993
FDPdur: FDP duration	X_5	-.371	.690	.545	.873
MinPassedTOD: Duty time passed at TOD	X_6	.000	1.000	1.000	1.000
TZ_WE: Time zones crossed from West to East	X_7	.000	1.000	1.000	1.000
SleepOpp48: Sleep opportunity in 48h prior TOD	X_8	-.001	.999	.998	.999
SleepOppD48: Sleep opportunity in darkness in 48h prior TOD	X_9	3.234	25.390	17.773	36.273
AwakeTOD: Time awake prior TOD	X_{10}	.222	1.248	1.222	1.275

OR Odds Ratio. CI Confidence Interval.

We see that finishing later, shorter FDP duration, more sleep opportunity in darkness in 48 hours prior TOD, and longer time awake show a relatively strong increase in odds of high fatigue. For the CAS measure there is no effect of duty time passed, time zones crossed, and sleep opportunity in 48 hours on the odds of high fatigue. The variable late finish overlaps with the grouping variable for consecutive late finishes which explains the high odds ratio.

<u>Consecutive FDP2 (Nights)</u>

This resulted in the inclusion of six independent variables for the CAS dependent variable and seven for the Effect dependent variable, and a constant term. The other independent variables were excluded by the regression method due to redundancies. Table 27 summarizes the results of this step.

For the CAS measure there is no effect of sleep on the odds of high fatigue, similar to the results for FDP1 (Night duties of more than 10 hours). Also for Effect we see only a weak increase in the odds for the sleep variables. Later end time of the duty period shows a relatively strong increase in odds for the CAS measure. The Effect measure shows a significant FDP start time and alternative WOCL variable. We see contradicting direction of ORs for (shorter/longer) flight/FDP duration between the models. Rest period and time awake show hardly an effect on the odds.

Table 27 Results of multiple regression method for consecutive FDP2 (Nights)

Variable *CAS TOD*	***Symbol***	***b***	***OR***	***95% CI for OR***	
				Lower	*Upper*
Constant	X_0	-11.689	.000		
SleepOppD24: Sleep opportunity in darkness in 24h prior TOD	X_1	.000	1.000	1.000	1.000
SleepOppD48: Sleep opportunity in darkness in 48h prior TOD	X_3	.000	1.000	1.000	1.000
SleepOpp24: Sleep opportunity in 24h prior TOD	X_4	.000	1.000	1.000	1.000
RestPeriod: Rest period	X_6	.002	1.002	1.001	1.004
FDPdur: FDP duration	X_7	-.013	.987	.982	.992
EndH: FDP end time	X_9	1.236	3.442	2.554	4.638
Variable *Effect TOD*	***Symbol***	***b***	***OR***	***95% CI for OR***	
				Lower	*Upper*
Constant	X_0	-1.485	.227		
FlightDur: Flight duration	X_1	-.013	.987	.984	.990
StartH: FDP start time	X_2	-.251	.778	.754	.803
SleepPred24: Sleep prediction in 24h prior TOD	X_4	.006	1.006	1.004	1.008
AwakeTOD: Time awake prior TOD	X_5	-.004	.996	.995	.997
FDPaltWOCLm: Time in alternative WOCL	X_6	-.025	.976	.971	.980
FDPdur: FDP duration	X_7	.011	1.011	1.009	1.014
SleepPred48: Sleep prediction in 48h prior TOD	X_{10}	.006	1.006	1.005	1.008

OR Odds Ratio. CI Confidence Interval.

FDP2 (Mix)

This resulted in the inclusion of eight independent variables for the CAS dependent variable and six for the Effect dependent variable, and a constant term. The other independent variables were excluded by the regression method due to redundancies. Table 28 summarizes the results of this step.

In neither model we see an effect of sleep on the odds of high fatigue. Later end time of the duty period shows a relatively strong increase in odds for the CAS measure. The variables early start and (very) late finish were included in the CAS model; these variables overlap with the grouping variable for consecutive mix. The same goes for the transitions from late finish to night duties in the Effect model. Several WOCL variables show an effect (some contradictory) on the odds of high fatigue. Later time of day increases the odds for high fatigue.

Table 28 Results of multiple regression method for FDP2 (Mix)

***Variable** CAS TOD*	***Symbol***	***b***	***OR***	***95% CI for OR***	
				Lower	*Upper*
Constant	X_0	-18.286	.000		
SleepOppD24: Sleep opportunity in darkness in 24h prior TOD	X_1	.000	1.000	1.000	1.000
StartH: FDP start time	X_2	.028	1.029	1.001	1.058
EndH: FDP end time	X_3	.510	1.666	1.524	1.821
FDP2ES: FDP2 early start	X_4	6.304	546.624	97.160	3075.322
FDP2LF: FDP2 late finish	X_6	8.122	3368.072	1306.502	8682.655
VeryLF: Very late finish	X_8	-3.349	-3.349	-3.349	-3.349
SleepOpp24: Sleep opportunity in 24h prior TOD	X_9	.000	1.000	1.000	1.000
ES: Early start	X_{10}	-2.709	.067	.010	.462
***Variable** Effect TOD*	***Symbol***	***b***	***OR***	***95% CI for OR***	
				Lower	*Upper*
Constant	X_0	3.487	32.700		
FDPaltWOCLm: Time in alternative WOCL	X_1	-.074	.929	.913	.945
FDPWOCLm: Time in WOCL	X_3	.025	1.026	1.010	1.041
FDPaltWOCL1h: At least 1h in alternative WOCL	X_4	-1.795	.166	.034	.823
LF2N: Transition from late finish to night	X_5	5.576	264.097	32.873	2121.748
FDPWOCL%: Percentage in WOCL	X_6	-.027	.974	.960	.987
TimeDayTOD: Time of day at TOD	X_9	.120	1.128	1.061	1.199

OR Odds Ratio. CI Confidence Interval.

Planned vs worked rosters

Paired datasets of planned and worked rosters from four airlines were compared in order to study the stability of the roster planning and its impact on level of fatigue. For the remaining two participating airlines we could not pair the planned and worked rosters due to a misalignment in periods for which we received their rosters, due to incompleteness, or due to format issues. The resulting dataset contained 207,318 paired records.

Arrival delays

Of the entire paired roster dataset more than 92% of flights arrived on-time or 30 minutes late at the most (i.e., based on the comparison made between planned and worked rosters). Figure 15 shows that long-haul flights had longer delays than the other haul types (15.4%). Next was medium-haul with 9.9% delays, short-haul with 7.4%, cargo with 6.6%, and the least arrival delays were for the regional flights (4.6%).

The data was further analysed to study how delays of more than 30 minute affect the values of Effect and CAS. We performed a repeated-measures Anova between planned and worked data per haul types (5 groups).

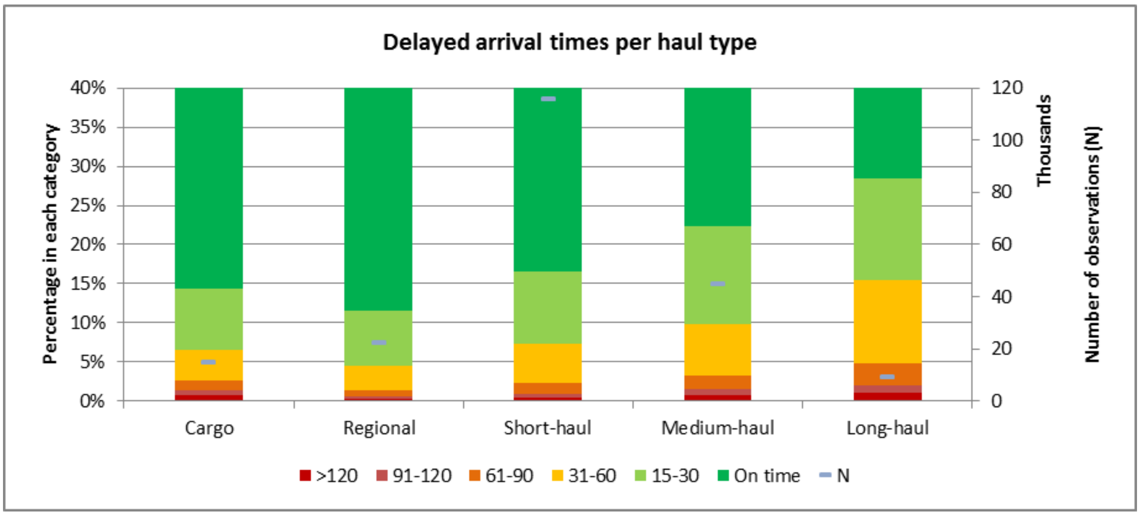

Figure 15 Delay in arrival times per haul type

Effect

The main effect of planned versus worked is significant ($F(1) = 99.041$, $p < .01$). However, this effect was very small (partial $\eta^2 = .003$). The main effect of haul type is also significant ($F(4) = 6240.117$). This effect was larger (partial $\eta^2 = .391$). Post-hoc analysis shows that the Effect scores were lowest for long haul ($M = 81.9$, $SE = .104$), followed by cargo ($M = 85.9$, $SE = .114$), medium haul ($M = 94.5$, $SE = .053$), regional ($M = 96.5$, $SE = .089$), and short haul ($M = 97.0$, $SE = .038$). All differences were significant ($p < .01$). *M* = mean and *SE* = standard error.

The interaction effect of roster and haul type was significant ($F(4, 38867) = 62.247$, $p < .01$). This effect was very small (partial $\eta^2 = .006$). All results are shown in Figure 16.

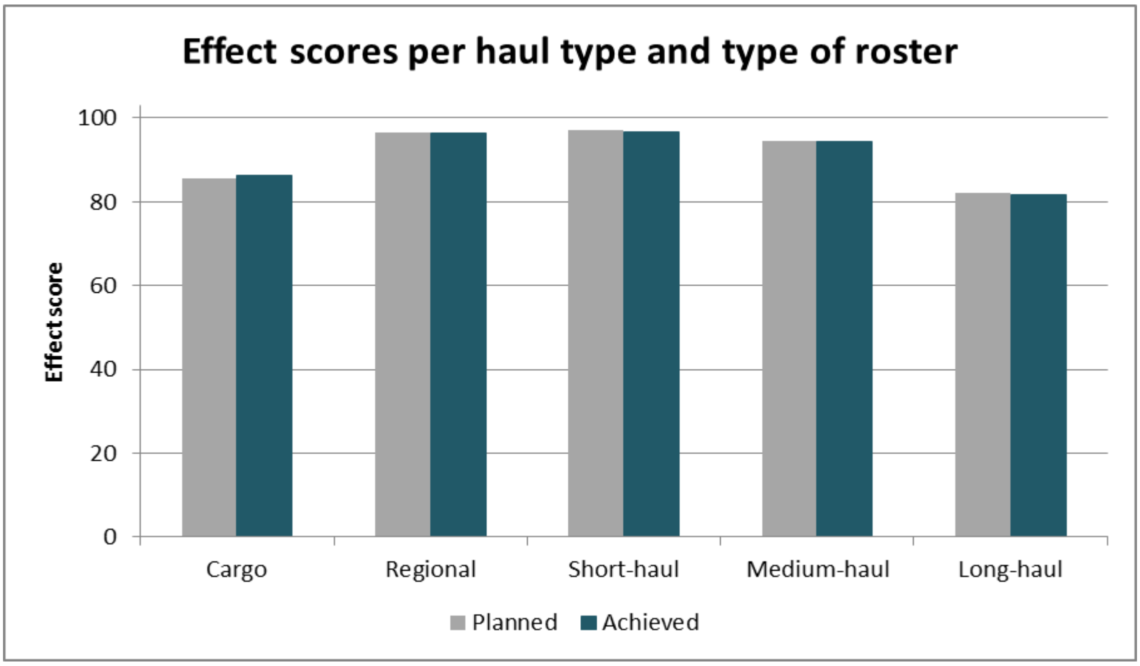

Figure 16 Mean Effect scores (with a range between 20 and 103) per haul type and type of roster for flights with delays of more than 30 minutes

CAS

Planned versus worked shows a significant main effect ($F(1) = 528.345$, $p < .01$). However, this effect was very small (partial $\eta^2 = .013$). The main effect of haul type is also significant ($F(4) = 80.989$), but very small (partial $\eta^2 = .018$). Post-hoc analysis shows that the Effect scores were lowest for long haul ($M = 2670.9$, $SE = 99.9$), followed by cargo ($M = 3266.1$, $SE = 110.1$), medium haul ($M = 4891.84$, $SE = 51.1$), regional ($M = 4891.9$, $SE = 86.0$), and short haul ($M = 5096.9$, $SE = 36.5$). All differences were significant ($p < .01$), except for the differences between regional, short haul and medium haul. M = mean and SE = standard error.

The interaction effect of roster and haul type was significant ($F(4, 38867) = 80.989$, $p < .01$). This effect was very small (partial $\eta^2 = .008$). All results are shown in Figure 17.

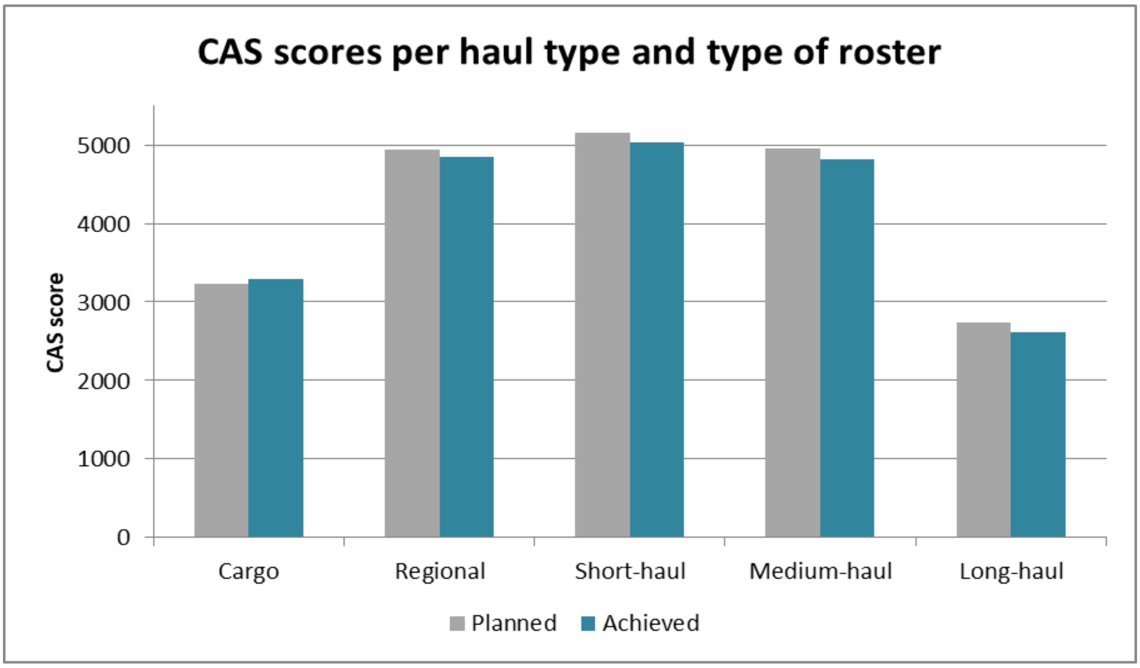

Figure 17 Mean CAS scores (with a range between 0 and 10,000) per haul type and type of roster for flights with delays of more than 30 minutes

Flight duration and sleep in the past 24 hours

The differences between planned flights and worked flights were compared for flight duration; predicted sleep (for Effect) or sleep opportunity (for CAS) in the past 24 hours. Each difference was calculated using a paired-samples t-test and all differences were significant (Table 29). However, the effect sizes (i.e., Cohen's d) are all very small[13]. Therefore, the significance could not be explained by large differences, but was likely caused by the relatively large sample size.

[13] A moderate effect size is around .5 and a small effect size around .2.

Table 29 Paired-samples t-test

Variable	*Degrees of freedom*	*t-value*	*p*	*Mean planned*	*Mean worked*	*Mean difference*	*Cohen's d*
Flight duration	207,317	159.097	< .01	130.92	127.44	3.482	.027
SleepOpp24	207,317	15.764	< .01	792.521	789.601	2.919	.014
SleepPred24	207,317	12.439	< .01	478.361	479.144	.782	.010

Flight duration, SleepOpp24 and SleepPred24 in minutes.

Percentage of high predicted fatigue scores

Figure 18 displays the percentage of high scores on Effect for each type of flight. It shows that over 46% of long-haul flights, between 12% and 13% of the cargo flights, and between 5% and 6% of medium-haul flights result in high predicted fatigue scores. Note that the differences between the planned and worked flights resulting in high fatigue scores are very small. Figure 19 shows a similar trend for the different haul types on CAS, except for long haul where the percentage of high predicted fatigue scores shows an increase of 5% from planned to worked flights.

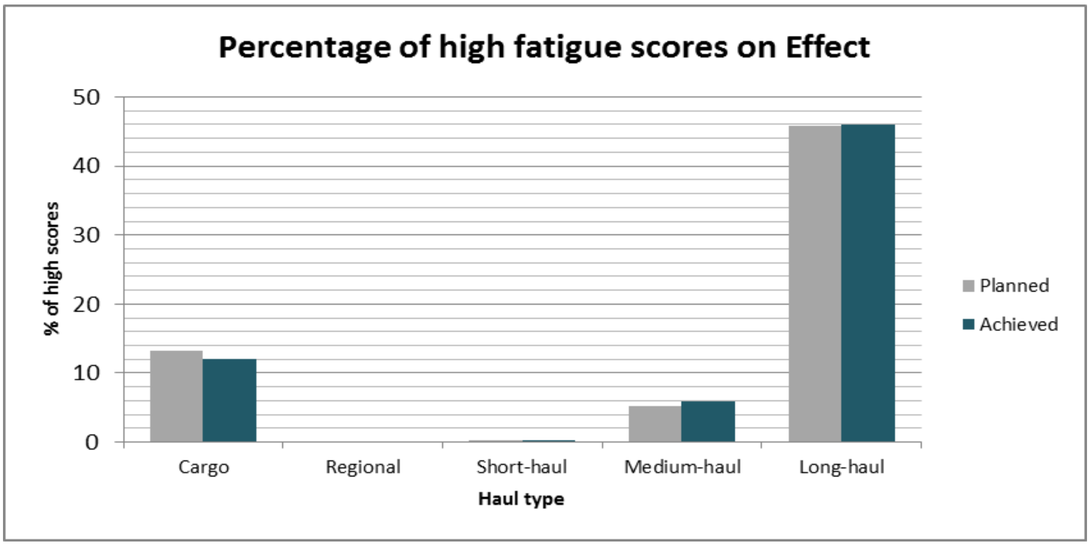

Figure 18 Percentage of high fatigue scores on Effect

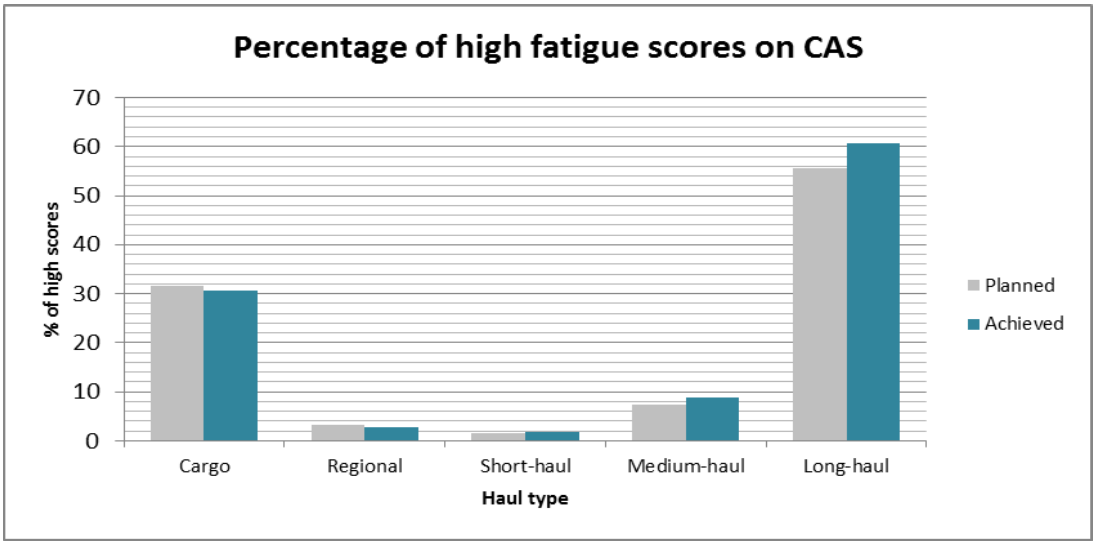

Figure 19 Percentage of high fatigue scores on CAS

CONCLUSION

The goal of this step was to check the stability of the roster planning and its possible impact on level of fatigue. Relevant to mention is that 15% of the survey respondents indicated 'delays' as a condition that may worsen fatigue. The item 'changes in the schedule' was indicated in 10% of the time.

With regard to the item 'delays' the roster analysis (using the current dataset with four paired airlines) did not result in any evidence for the level of fatigue being impacted by delayed arrivals.

The survey results on the item 'changes in the schedule' could not be verified in this particular roster analysis because of the limitations on the datasets. There was no way to determine on an individual crew member level whether or not a (late) change in flight schedule had a worsening effect on predicted sleep and level of fatigue.

Chapter 5: Mapping the identified fatigue hotspots

Roster data representativeness

The airlines that shared their rosters were used as a reference set for the EU aviation sector. The data analysis was performed with this set. Looking at the geographical distribution and type of operations that was included in the set, the dataset lacked rosters from the Northern region of Europe[14].

It turned out to be quite a time-consuming effort to retrieve the worked and planned rosters. Especially for the planned rosters airlines struggled to share the rosters before the deadline that was given to them by the project team. We were therefore not able to retrieve the planned rosters for all participating airlines.

Rosters were collected for a period of twelve months. During these months both low- and high-activity periods were covered. The months of July to October showed the highest numbers of FDPs in the current roster dataset.

FDP1 (Night duties of more than 10 hours)

Check for high predicted fatigue scores

We sought to assess the prevalence of high predicted fatigue scores during duties of more than 10 hours encroaching on the night. Our results show that – according to the bio-mathematical models used to examine the roster data – high fatigue scores did occur in flight duties longer than 10 hours that encroached partially or fully on the period between 02:00h and 04:59h. Moreover, the proportion of high fatigue was greater in these FDPs than in the baseline set containing all the collected FDPs. This was confirmed by the survey results, as 14% of respondents indicated 'a long working day' and 11% of respondents indicated 'flying during hours when I would normally sleep' as fatigue factors, and both were linked to night duties of more than 10 hours.

Find clusters of variables

We sought clusters of FDP-related characteristics that might impact the level of predicted fatigue during night FDPs of more than 10 hours. We worked our way towards multiple logistic regression models for the CAS and Effect measures, estimated for TOD during the final sector of the FDP.

The resulting multiple regression models (based on either the CAS or Effect dependent variable) differed to some extent, which can only be the result of differences between the bio-mathematical models used, as the same datasets were used in the analysis. Although both models are 'two-process models', they are not alike in how they represent and implement the two processes; i.e., the mathematical representation of each process is different, the relative weighting of the two processes is different, and the manner in which the two models estimate a pattern of sleep associated with a sequence of duties is different. The two models trace their roots back to different research data (either primarily validated against alertness ratings or cognitive performance). This difference may account for the differences in weightings of fatigue factors.

[14] Check D2.2 (Definition of the Data Collection Process) for an estimation of the size and geographical distribution of the entire aircrew population base.

The resulting FDP1 multiple regression models (also referred to as clusters of variables) included the following predictors:

- Sleep prediction in 24h prior TOD of the final sector;
- Start and end time of the FDP;
- Time in (alternative) window of circadian low (WOCL: 02:00h - 05:59h); and
- Number of time zones crossed east- and westwards.

The predicted duration of sleep in the 24 hours preceding the FDP was included in the multiple regression model because sleep is the primary recovery mechanism for fatigue. The bio-mathematical models differ in the way they estimate sleep; e.g., CAS includes no consideration of the effect of sleep on the odds of high fatigue. There is, however, an indirect link, as 23% of survey respondents indicated 'insufficient time between duties' as one of the contributing factors to fatigue. Insufficient time between duties may be interpreted as not enough time to get a good sleep.

For night FDPs of more than 10 hours, the earlier start and later end time of the duty period were included in the multiple regression models. This was not surprising, as the body's circadian rhythm has a major effect on fatigue levels. This is confirmed by the significance of the WOCL variable. Here, we see a clear link with the survey results, as 39% of respondents indicated that they were most fatigued 'in the WOCL', and 22% replied 'at the end of the duty'. Nine percent of the respondents indicated 'flying or being awake during hours when I would normally sleep' as a factor contributing to fatigue. Note that in the survey fatigue was not reported specifically at TOD, as opposed to in the roster data analyses.

For night FDPs of more than 10 hours the number of time zone crossings was included in the models because crossing more time zones results in longer FPDs that potentially encroach (part of) the WOCL. There was only a limited reference to time zone crossings in the survey responses; i.e., in- and outbound flights crossing more than six time zones was indicated in around 3% of the cases.

<u>Relevant rules in Subpart FTL</u>

The predictors included in the multiple regression model for flight duties longer than 10 hours that encroached partially or fully on the night are linked directly or indirectly to the following rules in the flight duty time limitations and rest requirements (Commission Regulation (EU) No. 83/2014):

- Start of the FDP at reference time (ORO.FTL.205 FDP);
- Sleep opportunity in 24h and 48h (ORO.FTL.205 FDP and ORO.FTL.235 Rest Periods);
- Flight duration (ORO.FTL.205 FDP and ORO.FTL.210 Flight Times and Duty Periods);
- Number of time zones crossed (ORO.FTL.235 Rest Period); and
- Duty time in WOCL (ORO.FTL.205 FDP).

FDP2 (Disruptive schedules)

Check for high predicted fatigue scores

We sought to assess the prevalence of high predicted fatigue scores during the different types of consecutive or non-consecutive disruptive flight duties. According to the bio-mathematical models used to examine the roster data, high predicted fatigue scores did occur for most types of disruptive schedules, but there were some differences between the different types of schedules.

For the *non-consecutive* disruptive schedules we found a relatively high prevalence of fatigue for the late finishes and nights. This was underpinned by the survey results, as 7% of respondents indicated 'late finishes' and 11% indicated 'flying during hours when I would normally sleep' as causes of fatigue. This was not the case for early starts, for which the prevalence of fatigue was low. In somewhat of a contradiction, a relatively high percentage of survey respondents (12%) indicated 'starting early' as a relevant fatigue item. However, it should be noted here that the survey questions used 'fatigue' as a broad term including physical fatigue, mental fatigue, and sleepiness and that the survey questions were not specifically focused on TOD. Nonetheless, the proportion of high fatigue was greater for the non-consecutive disruptive duties than for the baseline set, except for early starts, and the proportion for late finishes was just below the 1.0 relative ratio for Effect.

For the *consecutive* disruptive schedules (i.e., at least two in a row) we found a relatively high prevalence of fatigue for nights. The two bio-mathematical models showed different outcomes for late finishes: high prevalence for CAS and very low prevalence for Effect. For early starts the prevalence of fatigue was very low. The proportion of high fatigue was greater for the consecutive night disruptive schedules than in the FDP baseline set. The proportion of high fatigue was (effectively) zero for the consecutive early starts and results were inconclusive for the consecutive late finishes; i.e., the proportion was larger than 1 for CAS and zero for Effect.
Note that the same FDP baseline set, with the same implications, was used for the disruptive flight schedules as for the night FDPs of more than 10 hours.

Find clusters of variables

We sought clusters of FDP-related characteristics that might impact the level of predicted fatigue during *consecutive* disruptive flight duties. Similar to the approach used for FDP1 (Night duties of more than 10 hours) we worked our way towards multiple logistic regression models for the CAS and Effect measure. The conclusions drawn below are based on the independent variables included in the resulting multiple regression models.

The resulting multiple regression models per consecutive disruptive duty differed to some extent which results from the differences between the bio-mathematical models.

<u>Consecutive early starts</u>
Multiple regression models for consecutive early starts could not be computed because the prevalence of high predicted fatigue was very low; i.e., a prevalence of 1 (CAS) and 0 (Effect) in 6,895 observations. The survey results indicated 'consecutive early starts' to describe how the preceding roster affected their fatigue in 4% of the cases.

Consecutive late finishes

Only for the CAS measure could a multiple logistic regression model be developed as there were no valid cases for modelling of the Effect measure. The resulting FDP2 model for consecutive late finishes included the following predictors:

- Late finishes;
- Sleep opportunity in 48h in darkness prior TOD of the final sector;
- FDP duration; and
- Time awake prior TOD.

The variable very late finish shows a relatively strong increase in odds of high fatigue. The survey results indicated 'finishing late' as a relevant fatigue item in about 7% of the cases.

The fact that duration of night-time sleep opportunity in the 48 hours preceding the FDP was included in the multiple regression model was expected, because sleep is the primary recovery mechanism for fatigue. There is an indirect link to the survey result that 33% indicated 'insufficient time between duties' as one of the contributing factors to fatigue.

With regard to FDP duration, too, the survey results provide a link. Some 8% of respondents indicated 'long working days' as a contributor to fatigue. This was confirmed in the model by the significance of an extended time awake prior to TOD. In addition, 32% of survey respondents indicated 'at the end of the duty' as answer to the question 'when do you feel most fatigued'.

Consecutive night flights

The resulting FDP2 multiple regression models for consecutive nights included to following predictors:

- Time in (alternative) WOCL;
- Start and end time of the FDP; and
- FDP duration.

Sleep opportunity (or prediction) did not emerge as a strong predictor in the multiple regression models. Another recovery-related predictor that was not included is rest period. This refers to the period directly prior to the FDP.
WOCL was identified as a predictor in the model for the Effect measure. The importance of this variable appeared in the survey results as well. In response to the question 'when do you feel most fatigued', 38% of respondents indicated 'in the WOCL' as one of their answers. Seven percent of the respondents indicated 'flying or being awake during hours when I would normally sleep' as a factor contributing to fatigue; and 8% indicated 'unfavourable time for resting' as a factor.

For consecutive night flights, a correlation with fatigue at TOD was found for the start and end of the FDP at the reference time. This is similar to the results for night FDPs of more than 10 hours and is aligned with previous studies as well. The physiological mechanism underlying this phenomenon is well known. Fatigue during night FDPs is due to the circadian downswing of alertness and extended time awake. In addition, the amount of night sleep may be short, especially before inbound night flights if the local time and the biological clock are misaligned, causing circadian disruption.

Survey respondents (8%) indicated 'a long working day' as a contributor to high fatigue in disruptive flight schedules. This variable (i.e., FDP/flight duration) was also included in the multiple regression models and seems to be related to the end and start hour of the FDP.

Mix of early starts, late finishes and night flights

The multiple regression models for a mix of disruptive schedules included the following predictors:

- Early starts and late finishes;
- Transitions from late finish to night;
- Time in WOCL;
- Start and end of FDP; and
- Time of day.

The variables early start and (very) late finish were included in the CAS model; these variables overlap with the grouping variable for mix of disruptive schedules. Given the grouping variable, it also makes sense that several WOCL variables were included as predictors. The same goes for the transitions from late finish to night duties.

Sleep was not identified as a predictor in either model. Later start and end time of the duty period and later time of day increase the odds for high fatigue.

Relevant rules in Subpart FTL

The relevant predictors for the different consecutive disruptive schedules are linked directly or indirectly to the following rules in the flight duty time limitations and rest requirements (Commission Regulation (EU) No. 83/2014):

- Start of the FDP at reference time (ORO.FTL.205 FDP);
- Sleep opportunity in 24h and 48h (ORO.FTL.205 FDP and ORO.FTL.235 Rest Periods);
- Flight duration (ORO.FTL.205 FDP and ORO.FTL.210 Flight Times and Duty Periods);
- Number of time zones crossed (ORO.FTL.235 Rest Period);
- Duty time in WOCL (ORO.FTL.205 FDP); and
- Rest period provided before undertaking an FDP (ORO.FTL.235 Rest Period).

List of abbreviations

Abbreviations	***Description***
Anova	Analysis of variance
BAM	Boeing Alertness Model
CAS	Common Alertness Scale
CAT	Commercial Air Transport
CI	Confidence Interval
D	Deliverable
EASA	European Aviation Safety Agency
EC	European Commission
EU	European Union
FAST	Fatigue Avoidance Scheduling Tool
FDP	Flight Duty Period
FTL	Flight Time Limitation
KSS	Karolinska Sleepiness Scale
M	Mean
Max	Maximum
OR	Odds Ratio
ORO	Organisation Requirements (in the air Operations Regulation)
SAFTE	Sleep, Activity, Fatigue, and Task Effectiveness
SE	Standard Error
SP	Samn-Perelli
TOD	Top Of Descent
WOCL	Window Of Circadian Low

Appendix 1: Online survey

Section A: Informed consent

A1.

The collected survey data will be separated from your identity. No information that you provide will be published that allows you as an individual to be identified. We certify to treat collected data according to good practice and follow sound ethical rules in accordance with Regulation (EC) No 45/2001 of the European Parliament and of the Council of 18 December 2000 on the protection of individuals with regard to the processing of personal data by the Community institutions and bodies and on the free movement of such data.

Please be aware that the answers you provide to this survey will be collected and analysed and that you can stop filling out questions at any time without the need to provide any explanation. Your input to this questionnaire is voluntary. If you have any questions or comments concerning this study you can contact:

Dr. H. van Dijk
Netherlands Aerospace Centre - NLREmail: info@nlr.nl

Do you understand and agree with the information stated above?

Yes ☐

Section B: Demographics

B1. **What is your occupation?**

Pilot ☐
Cabin crew ☐
Researcher ☐
Safety expert ☐
Other ☐

Other

B2. **What is your function?**

- Captain ☐
- First officer ☐
- Other ☐

Other

B3. **What is your function?**

- Chief purser ☐
- Purser ☐
- Flight attendant ☐
- Other ☐

Other

B4. **Do you currently work for an airline?**

- Yes ☐
- No ☐

B5. **For which type of airline do you currently work?**

- Leisure carrier ☐
- Network operator ☐
- Cargo ☐
- Other ☐

Other

B6. In which country is your airline based?

Please note that this questionnaire is intended for employees who work for European airlines only.

- Albania ☐
- Andorra ☐
- Armenia ☐
- Austria ☐
- Azerbaijan ☐
- Belgium ☐
- Bosnia and Herzegovina ☐
- Bulgaria ☐
- Croatia ☐
- Cyprus ☐
- Czech Republic ☐
- Denmark ☐
- Estonia ☐
- Finland ☐
- France ☐
- Georgia ☐
- Germany ☐
- Greece ☐
- Hungary ☐
- Iceland ☐
- Ireland ☐
- Italy ☐
- Kosovo ☐
- Latvia ☐
- Liechtenstein ☐
- Lithuania ☐
- Luxembourg ☐
- Macedonia ☐
- Malta ☐

B7. What is your country of residence?

- [ ] Albania
- [ ] Andorra
- [ ] Armenia
- [ ] Austria
- [ ] Azerbaijan
- [ ] Belgium
- [ ] Bosnia and Herzegovina
- [ ] Bulgaria
- [ ] Croatia
- [ ] Cyprus
- [ ] Czech Republic
- [ ] Denmark
- [ ] Estonia
- [ ] Finland
- [ ] France
- [ ] Georgia
- [ ] Germany
- [ ] Greece
- [ ] Hungary
- [ ] Iceland
- [ ] Ireland
- [ ] Italy
- [ ] Kosovo
- [ ] Latvia
- [ ] Liechtenstein
- [ ] Lithuania
- [ ] Luxembourg
- [ ] Macedonia
- [ ] Malta
- [ ] Moldova
- [ ] Monaco

B8. **What is the travel time from your residence to work (please use whole numbers only)?**

Hours

Minutes

Section C: Schedule

C1. **How is the current distribution of the flights that you make in percentages (please make sure that the added value is 100%)?**

Regional (< 1 hour)

Short-haul (1-2 hours)

Medium-haul (> 2 hours)

Long-haul (> 5 hours and crossing > 2 time zones)

C2. **What was the distribution of the flights that you made when working for an airline in percentages (please make sure that the added value is 100%)?**

Regional (< 1 hour)

Short-haul (1-2 hours)

Medium-haul (> 2 hours)

Long-haul (> 5 hours and crossing > 2 time zones)

C3. **Think about the last time you were in active duty (not on a positioning flight) and experienced fatigue you believe was caused by the schedule. Then think about your schedule the days before feeling so fatigued. Please indicate from the list below one or more items that you deem relevant for causing the situation.**

- Starting early ☐
- Finishing late ☐
- Long working days ☐
- Not sleeping at home several nights in a row ☐
- Outward westward flight across >6 time zones ☐
- Outward eastward flight across >6 time zones ☐
- Return flight after a westward flight across >6 time zone ☐
- Return flight after an eastward flight across >6 time zones. ☐
- Flying a great number of sectors ☐

Unfavourable times for resting (e.g. rest period when you are not sleepy) ☐

Short recovery time between duties ☐

Insufficient rest time during flight ☐

Insufficient quality of on-board rest facilities ☐

Flying during hours when I would normally sleep ☐

Other ☐

Other

C4. Please explain about the nature of your roster and why this combination of items resulted in you being so fatigued?

Section D: Flight Duty Periods

D1. The following questions concern the following 6 Flight Duty Periods (see below for a description of each Flight Duty Period):

Duties of more than 13 hours at the most favourable time of the day (this refers to daytime operations (from 8:00h to 22:00h)). Duties of more than 10 hours at less favourable time of the day (this refers to operations that encroach (part of) the night (the period between 02:00h and 05:00h)). Duties of more than 11 hours for crewmembers in an unknown state of acclimatisation* Duties including a high level of sectors (more than 6; this refers to daytime operations (from 06:00h to 22:00h)). On-call duties such as standby or reserve followed by flight duties (this refers to daytime flight duties). Disruptive schedules**

Please indicate which Flight Duty Periods you have experienced in the past three years.

** "Acclimatised" means a state in which a crew member's biological clock is synchronised to the time zone where the crew member is. A crew member is considered to be acclimatised to a 2-hour wide time zone surrounding the local time at the point of departure. When this local time differs by more than 4 hours from the local time at the place where the next duty starts, the crew member can get in an unknown state of acclimatisation, dependent on the time difference. It takes from 48 up to 120 hours to get acclimatised again. Before 48 hours you are considered to still be acclimatised to the local time at the point of departure.*

*** A disruptive schedule means a crew member's roster which disrupts the sleep opportunity during the optimal sleep time window by comprising a duty or a combination of duties which encroach, start or finish, during any portion of the day or of the night where a crew member is acclimatised. A schedule may be disruptive due to early starts, late finishes or night duties (e.g. "early type" of disruptive schedule means: for "early start" a duty period starting in the period between 05:00h and 06:00h in the time zone to which a crew member is acclimatised; and for "late finish" a duty period finishing in the period between 23:00h and 02:00h; "late type" of disruptive schedule means for "early start" a duty period starting in the period between 05:00h and 07:00h in the time zone to which a crew member is acclimatised; and for "late finish" a duty period finishing in the period between 00:00h and 02:00h.*

Please check the COMMISSION REGULATION (EU) No 83/2014 for more detailed definitions of these duties.

- ☐ Duties of more than 13 hours at the most favourable time of the day
- ☐ Duties of more than 10 hours at less favourable time of the day
- ☐ Duties of more than 11 hours for crewmembers in an unknown state of acclimatisation
- ☐ Duties including a high level of sectors (more than 6)
- ☐ On-call duties such as standby or reserve followed by flight duties
- ☐ Disruptive schedules
- ☐ None of the above

D2. The following questions concern the following 6 Flight Duty Periods (see below for a description of each Flight Duty Period):

Duties of more than 13 hours at the most favourable time of the day (this refers to daytime operations (from 8:00h to 22:00h)). Duties of more than 10 hours at less favourable time of the day (this refers to operations that encroach (part of) the night (the period between 02:00h and 05:00h)). Duties of more than 11 hours for crewmembers in an unknown state of acclimatisation* Duties including a high level of sectors (more than 6; this refers to daytime operations (from 06:00h to 22:00h)). On-call duties such as standby or reserve followed by flight duties (this refers to daytime flight duties). Disruptive schedules**

Please indicate which Flight Duty Periods exist in your airline's roster.

** "Acclimatised" means a state in which a crew member's biological clock is synchronised to the time zone where the crew member is. A crew member is considered to be acclimatised to a 2-hour wide time zone surrounding the local time at the point of departure. When this local time differs by more than 4 hours from the local time at the place where the next duty starts, the crew member can get in an unknown state of acclimatisation, dependent on the time difference. It takes from 48 up to 120 hours to get acclimatised again. Before 48 hours you are considered to still be acclimatised to the local time at the point of departure.*

*** A disruptive schedule means a crew member's roster which disrupts the sleep opportunity during the optimal sleep time window by comprising a duty or a combination of duties which encroach, start or finish, during any portion of the day or of the night where a crew member is acclimatised. A schedule may be disruptive due to early starts, late finishes or night duties (e.g. "early type" of disruptive schedule means: for "early start" a duty period starting in the period between 05:00h and 06:00h in the time zone to which a crew member is acclimatised; and for "late finish" a duty period finishing in the period between 23:00h and 02:00h; "late type" of disruptive schedule means for "early start" a duty period starting in the period between 05:00h and 07:00h in the time zone to which a crew member is acclimatised; and for "late finish" a duty period finishing in the period between 00:00h and 02:00h.*

Please check the COMMISSION REGULATION (EU) No 83/2014 for more detailed definitions of these duties.

- ☐ Duties of more than 13 hours at the most favourable time of the day
- ☐ Duties of more than 10 hours at the less favourable time of the day
- ☐ Duties of more than 11 hours for crewmembers in an unknown state of acclimatisation
- ☐ Duties including a high level of sectors (more than 6)
- ☐ On-call duties such as standby or reserve followed by flight duties
- ☐ Disruptive schedules
- ☐ None of the above

D3. **How many times per month do you typically meet the Flight Duty Periods?**

Duties of more than 13 hours at the most favourable time of the day

Duties of more than 10 hours at the least favourable time of the day

Duties of more than 11 hours for crewmembers in an unknown state of acclimatisation

Duties including a high level of sectors (more than 6)

On-call duties such as standby or reserve followed by flight duties

Disruptive schedules

Section E: Duties of more than 13 hours at the most favourable time of the day

E1. **The following questions are about Flight Duty Periods of more than 13 hours at the most favourable time of the day.**

Please describe how the roster preceding such duties affects your fatigue.

E2. **Concerning this Flight Duty Period, when do you feel most fatigued and what makes this type of duty so fatiguing?**

E3. **Please indicate how you would rate your alertness during the most fatiguing part of this type of Flight Duty Period.**

1. Extremely alert ☐
2. ☐
3. Alert ☐
4. ☐
5. Neither alert nor sleepy ☐
6. ☐
7. Sleepy - but no difficulty remaining awake ☐
8. ☐
9. Extremely sleepy - fighting sleep ☐

E4. **Concerning this Flight Duty Period, when do you think the aircrew feels most fatigued and what makes this type of duty so fatiguing?**

E5. **Which conditions may worsen fatigue in a Flight Duty Period of more than 13 hours at the most favourable time of the day?**

E6. **Which conditions may mitigate fatigue in a Flight Duty Period of more than 13 hours at the most favourable time of the day?**

E7. **Please indicate how often you typically encounter this type of Flight Duty Period.**

- Several times a week
- Between once a week and once a month
- Between once a month and once semi-anually
- Between once semi-anually and once anually
- Other

Other

E8. **Are duties of more than 13 hours at the most favourable time of the day typically flown with an augmented or non-augmented crew?**

- Augmented
- Non-augmented
- Other

Other

E9. **Please describe the typical crew composition during such duties.**

E10. Please indicate if in-flight rest facilities are typically available during duties of more than 13 hours at the most favourable time of the day.

'Class 1 rest facility' means a bunk or other surface that allows for a flat or near flat sleeping position. It reclines to at least 80° back angle to the vertical and is located separately from both the flight crew compartment and the passenger cabin in an area that allows the crew member to control light, and provides isolation from noise and disturbance;

'Class 2 rest facility' means a seat in an aircraft cabin that reclines at least 45° back angle to the vertical, has at least a pitch of 55 inches (137.5 cm), a seat width of at least 20 inches (50 cm) and provides leg and foot support. It is separated from passengers by at least a curtain to provide darkness and some sound mitigation, and is reasonably free from disturbance by passengers or crew members;

'Class 3 rest facility' means a seat in an aircraft cabin or flight crew compartment that reclines at least 40° from the vertical, provides leg and foot support and is separated from passengers by at least a curtain to provide darkness and some sound mitigation, and is not adjacent to any seat occupied by passengers.

- Yes, class 1 rest facility is available ☐
- Yes, class 2 rest facility is available ☐
- Yes, class 3 rest facility is available ☐
- No ☐
- Other ☐

Other

E11. Please indicate how much time during the flight you typically use the rest facility in Flight Duty Periods longer than 13 hours (please use whole numbers only).

Hours ____________

Minutes ____________

E12. Please indicate how much time during the flight the rest facility is typically used by each crew member in this type of Flight Duty Period.

Section F: Duties of more than 10 hours at less favourable time of the day

F1. **The following questions are about Flight Duty Periods of more than 10 hours at less favourable time of the day.**

Please describe how the roster preceding such duties affects your fatigue.

F2. **Concerning this Flight Duty Period, when do you feel most fatigued and what makes this type of duty so fatiguing?**

F3. **Please indicate how you would rate your alertness during the most fatiguing part of this type of Flight Duty Period.**

1. Extremely alert ☐
2. ☐
3. Alert ☐
4. ☐
5. Neither alert nor sleepy ☐
6. ☐
7. Sleepy - but no difficulty remaining awake ☐
8. ☐
9. Extremely sleepy - fighting sleep ☐

F4. **Concerning the Flight Duty Period of more than 10 hours at less favourable time of the day, when do you think the aircrew feels most fatigued and what makes this type of duty so fatiguing?**

F5. **Which conditions may worsen fatigue in a Flight Duty Period of more than 10 hours at less favourable time of the day?**

F6. **Which conditions may mitigate fatigue in a Flight Duty Period of more than 10 hours at less favourable time of the day?**

F7. **Please indicate how often you typically encounter this type of Flight Duty Period.**

- Several times a week ☐
- Between once a week and once a month ☐
- Between once a month and once semi-anually ☐
- Between once semi-anually and once anually ☐
- Other ☐

Other

F8. **Are duties of more than 10 hours at less favourable time of the day typically flown with an augmented or non-augmented crew?**

Augmented ☐

Non-augmented ☐

Other ☐

Other

F9. **Please describe the typical crew composition during such duties.**

F10. **Please indicate if in-flight rest facilities are typically available during duties of more than 10 hours at less favourable time of the day.**

'Class 1 rest facility' means a bunk or other surface that allows for a flat or near flat sleeping position. It reclines to at least 80° back angle to the vertical and is located separately from both the flight crew compartment and the passenger cabin in an area that allows the crew member to control light, and provides isolation from noise and disturbance;

'Class 2 rest facility' means a seat in an aircraft cabin that reclines at least 45° back angle to the vertical, has at least a pitch of 55 inches (137.5 cm), a seat width of at least 20 inches (50 cm) and provides leg and foot support. It is separated from passengers by at least a curtain to provide darkness and some sound mitigation, and is reasonably free from disturbance by passengers or crew members;

'Class 3 rest facility' means a seat in an aircraft cabin or flight crew compartment that reclines at least 40° from the vertical, provides leg and foot support and is separated from passengers by at least a curtain to provide darkness and some sound mitigation, and is not adjacent to any seat occupied by passengers.

Yes, class 1 rest facility is available ☐

Yes, class 2 rest facility is available ☐

Yes, class 3 rest facility is available ☐

No ☐

Other ☐

Other

F11. **Please indicate how much time during the flight you typically use the rest facility in a Flight Duty Period of more than 10 hours (please use whole numbers only).**

Hours

Minutes

F12. **Please indicate how much time during the flight the rest facility is typically used by each crew member in a Flight Duty Period of more than 10 hours.**

Section G: Duties of more than 11 hours for crew members in an unknown state of acclimatisation

G1. **The following questions are about Flight Duty Periods of more than 11 hours for crew members in an unknown state of acclimatisation.**

Please describe how the roster preceding such duties affects your fatigue.

G2. **Concerning this Flight Duty Period, when do you feel most fatigued and what makes this type of duty so fatiguing?**

G3. **Please indicate how you would rate your alertness during the most fatiguing part of this type of Flight Duty Period.**

1. Extremely alert ☐
2. ☐
3. Alert ☐
4. ☐
5. Neither alert nor sleepy ☐
6. ☐
7. Sleepy - but no difficulty remaining awake ☐
8. ☐
9. Extremely sleepy - fighting sleep ☐

G4. **Concerning the Flight Duty Period of more than more than 11 hours for crew members in an unknown state of acclimatisation, when do you think the aircrew feels most fatigued and what makes this type of duty so fatiguing?**

G5. **Which conditions may worsen fatigue in a Flight Duty Period of more than 11 hours for crew members in an unknown state of acclimatisation?**

G6. **Which conditions may mitigate fatigue in a Flight Duty Period of more than 11 hours for crew members in an unknown state of acclimatisation?**

G7. **Please indicate how often you typically encounter this type of Flight Duty Period.**

- ☐ Several times a week
- ☐ Between once a week and once a month
- ☐ Between once a month and once semi-anually
- ☐ Between once semi-anually and once anually
- ☐ Other

Other

G8. Are duties of more than 11 hours for crew members in an unknown state of acclimatisation typically flown with an augmented or non-augmented crew?

Augmented ☐

Non-augmented ☐

Other ☐

Other

G9. Please describe the typical crew composition during such duties.

G10. Please indicate if in-flight rest facilities are typically available on duties of more than 11 hours for crew members in an unknown state of acclimatisation.

'Class 1 rest facility' means a bunk or other surface that allows for a flat or near flat sleeping position. It reclines to at least 80° back angle to the vertical and is located separately from both the flight crew compartment and the passenger cabin in an area that allows the crew member to control light, and provides isolation from noise and disturbance;

'Class 2 rest facility' means a seat in an aircraft cabin that reclines at least 45° back angle to the vertical, has at least a pitch of 55 inches (137.5 cm), a seat width of at least 20 inches (50 cm) and provides leg and foot support. It is separated from passengers by at least a curtain to provide darkness and some sound mitigation, and is reasonably free from disturbance by passengers or crew members;

'Class 3 rest facility' means a seat in an aircraft cabin or flight crew compartment that reclines at least 40° from the vertical, provides leg and foot support and is separated from passengers by at least a curtain to provide darkness and some sound mitigation, and is not adjacent to any seat occupied by passengers.

Yes, class 1 rest facility is available ☐

Yes, class 2 rest facility is available ☐

Yes, class 3 rest facility is available ☐

No ☐

Other ☐

Other

G11. **Please indicate how much time during the flight you typically use the rest facility in a Flight Duty Period of more than 11 hours (please use whole numbers only).**

Hours

Minutes

G12. **Please indicate how much time during the flight the rest facility is typically used by each crew member in a Flight Duty Period of more than 10 hours.**

Section H: Duties including a high level of sectors (more than 6)

H1. **The following questions are about Flight Duty Periods including a high level of sectors (more than 6).**

Please describe how the roster preceding such duties affects your fatigue.

H2. **Concerning this Flight Duty Period, when do you feel most fatigued and what makes this type of duty so fatiguing?**

H3. **Please indicate how you would rate your alertness during the most fatiguing part of this type of Flight Duty Period.**

1. Extremely alert ☐
2. ☐
3. Alert ☐
4. ☐
5. Neither alert nor sleepy ☐
6. ☐
7. Sleepy - but no difficulty remaining awake ☐
8. ☐
9. Extremely sleepy - fighting sleep ☐

H4. **Concerning the Flight Duty Period of duties including a high level of sectors (more than 6), when do you think the aircrew feels most fatigued and what makes this type of duty so fatiguing?**

H5. **Which conditions may worsen fatigue in a Flight Duty Period of duties including a high level of sectors (more than 6)?**

H6. **Which conditions may mitigate fatigue in a Flight Duty Period of duties including a high level of sectors (more than 6)?**

H7. **Please indicate how often you typically encounter this type of Flight Duty Period.**

- Several times a week
- Between once a week and once a month
- Between once a month and once semi-anually
- Between once semi-anually and once anually
- Other

Other

H8. **Are duties including a high level of sectors (more than 6) typically flown with an augmented or non-augmented crew?**

- Augmented
- Non-augmented
- Other

Other

H9. **Please describe the typical crew composition during such duties.**

H10. Please indicate if in-flight rest facilities are typically available on duties including a high level of sectors (more than 6).

'Class 1 rest facility' means a bunk or other surface that allows for a flat or near flat sleeping position. It reclines to at least 80° back angle to the vertical and is located separately from both the flight crew compartment and the passenger cabin in an area that allows the crew member to control light, and provides isolation from noise and disturbance;

'Class 2 rest facility' means a seat in an aircraft cabin that reclines at least 45° back angle to the vertical, has at least a pitch of 55 inches (137.5 cm), a seat width of at least 20 inches (50 cm) and provides leg and foot support. It is separated from passengers by at least a curtain to provide darkness and some sound mitigation, and is reasonably free from disturbance by passengers or crew members;

'Class 3 rest facility' means a seat in an aircraft cabin or flight crew compartment that reclines at least 40° from the vertical, provides leg and foot support and is separated from passengers by at least a curtain to provide darkness and some sound mitigation, and is not adjacent to any seat occupied by passengers.

Yes, class 1 rest facility is available ☐

Yes, class 2 rest facility is available ☐

Yes, class 3 rest facility is available ☐

No ☐

Other ☐

Other

H11. Please indicate how much time during the flight you typically use the rest facility in Flight Duty Periods including more than 6 sectors (please use whole numbers only).

Hours

Minutes

H12. Please indicate how much time during the flight the rest facility is typically used by each crew member in Flight Duty Periods including more than 6 sectors.

Section I: On-call duties such as standby or reserve followed by flight duties

I1. **The following questions are about Flight Duty Periods of on-call duties such as standby or reserve followed by flight duties.**

Please describe how the roster preceding such duties affects your fatigue.

I2. **Concerning this Flight Duty Period, when do you feel most fatigued and what makes this type of duty so fatiguing?**

I3. **Please indicate how you would rate your alertness during the most fatiguing part of this type of Flight Duty Period.**

1. Extremely alert ☐
2. ☐
3. Alert ☐
4. ☐
5. Neither alert nor sleepy ☐
6. ☐
7. Sleepy - but no difficulty remaining awake ☐
8. ☐
9. Extremely sleepy - fighting sleep ☐

I4. **Concerning the Flight Duty Period of on-call duties such as standby or reserve followed by flight duties, when do you think the aircrew feels most fatigued and what makes this type of duty so fatiguing?**

I5. **Which conditions may worsen fatigue in on-call duties such as standby or reserve followed by flight duties?**

I6. **Which conditions may mitigate fatigue in on-call duties such as standby or reserve followed by flight duties?**

I7. **Please indicate how often you typically encounter this type of Flight Duty Period.**

- Several times a week ☐
- Between once a week and once a month ☐
- Between once a month and once semi-anually ☐
- Between once semi-anually and once anually ☐
- Other ☐

Other

Section J: Disruptive schedules

J1. **The following questions are about Flight Duty Periods of disruptive schedules.**

Please describe how the roster preceding such duties affects your fatigue.

J2. **Concerning this Flight Duty Period, when do you feel most fatigued and what makes this type of duty so fatiguing?**

J3. **Please indicate how you would rate your alertness during the most fatiguing part of this type of Flight Duty Period.**

1. Extremely alert ☐
2. ☐
3. Alert ☐
4. ☐
5. Neither alert nor sleepy ☐
6. ☐
7. Sleepy - but no difficulty remaining awake ☐
8. ☐
9. Extremely sleepy - fighting sleep ☐

J4. **Concerning the Flight Duty Period of on-call duties such as disruptive schedules, when do you think the aircrew feels most fatigued and what makes this type of duty so fatiguing?**

J5. **Which conditions may worsen fatigue in disruptive schedules?**

J6. **Which conditions may mitigate fatigue in disruptive schedules?**

J7. **Please indicate how often you typically encounter this type of Flight Duty Period.**

- Several times a week ☐
- Between once a week and once a month ☐
- Between once a month and once semi-anually ☐
- Between once semi-anually and once anually ☐
- Other ☐

Other

Section K: Ranking

K1. Please rank the following Flight Duty Periods on expected fatigue, even if you have no experience with them. Place the highest fatigue level on top and the lowest on the bottom. The Flight Duty Periods are further explained below.

Duties of more than 13 hours at the most favourable time of the day refers to daytime operations (from 8:00h to 22:00h). Duties of more than 10 hours at less favourable time of the day refers to operations that encroach (part of) the night (the period between 02:00h and 05:00h). Duties of more than 11 hours for crewmembers in an unknown state of acclimatisation: "Acclimatised" means a state in which a crew member's biological clock is synchronised to the time zone where the crew member is. A crew member is considered to be acclimatised to a 2-hour wide time zone surrounding the local time at the point of departure. When this local time differs by more than 4 hours from the local time at the place where the next duty starts, the crew member can get in an unknown state of acclimatisation, dependent on the time difference. It takes from 48 up to 120 hours to get acclimatised again. Before 48 hours you are considered to still be acclimatised to the local time at the point of departure. Duties including a high level of sectors (more than 6) refers to daytime operations (from 06:00h to 22:00h). On-call duties such as standby or reserve followed by flight duties refers to daytime flight duties. Disruptive schedules: A disruptive schedule means a crew member's roster which disrupts the sleep opportunity during the optimal sleep time window by comprising a duty or a combination of duties which encroach, start or finish, during any portion of the day or of the night where a crew member is acclimatised. A schedule may be disruptive due to early starts, late finishes or night duties (e.g. "early type" of disruptive schedule means: for "early start" a duty period starting in the period between 05:00h and 06:00h in the time zone to which a crew member is acclimatised; and for "late finish" a duty period finishing in the period between 23:00h and 02:00h; "late type" of disruptive schedule means for "early start" a duty period starting in the period between 05:00h and 07:00h in the time zone to which a crew member is acclimatised; and for "late finish" a duty period finishing in the period between 00:00h and 02:00h.

Please check the COMMISSION REGULATION (EU) No 83/2014 for more detailed definitions of these duties.

Duties of more than 13 hours at the most favourable time of the day ☐

Duties of more than 10 hours at less favourable time of the day ☐

Duties of more than 11 hours for crewmembers in an unknown state of acclimatisation ☐

Duties including a high level of sectors (more than 6) ☐

On-call duties such as standby or reserve followed by flight duties ☐

Disruptive schedules ☐

K2. Please elaborate on why you selected this ranking.

Section L: Conclusion

L1. For how many hours per week are you employed?

L2. **What type of contract do you have with your company?**

- Fixed-term ☐
- Permanent ☐
- Other ☐

Other

L3. **Do you perform an additional job within your company?**

- Yes ☐
- No ☐

L4. **What is your age in years?**

L5. **What is your gender?**

- Female ☐
- Male ☐
- Other ☐

Other

L6. **Do you have any comments, remarks or questions about this survey, the topic of this survey or the research? If yes, please type them here.**

L7. **Finally, please describe the event when you were most fatigued in your entire career. Try to provide as many details as possible: type of flight, flight phase, time of day, flight schedule in the days preceding the flight, non work related factors that might have affected your fatigue level, etc.**

Thank you for your participation!!!

Appendix 2: Model parameter settings

BAM parameter	***Description***	***Setting***
Briefing	Briefing time before the activity departure that should be seen as on-duty time in min	45
Debriefing	Debriefing after the activity that should be seen as on-duty time in min. Must not overlap the briefing on the following leg	30
Wake before	Enforced wakefulness before the leg in min (counting from start of briefing)	120
Wake after	Enforced wakefulness after the leg in min (counting from end of de-briefing)	180
Nap	Start and end time of in-flight nap in min. Use -1 if no nap	-1
SAFTE model parameter	***Description***	***Setting***
Duty break type	Selection from the following duty break types: **Sit:** Any gaps in between flights will be filled in with non-crewing sit time **Split Duty Sleep:** Any gaps between events will be filled with commute and/or sleep events **Open Time:** Any gaps between will be left events empty. No non-crewing or sleep events will be added. The auto-sleep algorithm will use that open time as if they are no longer on duty	Sit
Min sit for rest facility/hotel	Specification of the minimum sit time to occur, prior to the potential of a sleep event being added at a rest facility or hotel **Note:** by default, this item is zero and will result in no sleep at a rest facility or hotel during a break between flights	0
Commute duration	The commute time in min entered will apply to every case whether the crew member is travelling from home, hotel or rest facility	60
Auto sleep	Allows for sleep periods to be added automatically	Enabled
Min sleep duration	Minimum amount of time in min for a sleep event	60
Max work day sleep	Maximum amount of sleep time in minutes that will occur on a work day	480
Max rest day sleep	Maximum amount of sleep time in minutes that will occur on a rest day	540
Awake zone	Time of day that a person will not usually attempt to sleep due to different environmental and circadian factors	13-19h local
Normal bed time	The setting for when to determine the time of day when crew members will go to bed, provided they are not working	23h
Auto-nap	Crews who have been awake for an extended time will take naps prior to duties that start late at night. In these cases, the system will insert naps: **Line 1:** If the amount of time is 480-600 min (8-10h), then insert a 60-min nap prior to commute to next event **Line 2:** If the amount of time is 601-720 min (10-12h), then insert a 90-min nap prior to commute to next event **Line 3:** If the amount of time is 721-960 min (12-16h), then insert a 2h nap prior to commute to next event **Line 4:** If the amount of time is greater than 960 min (16h), then insert a 3h nap prior to commute to next event	See lines 1-4
Quality of sleep	**Excellent:** This environment setting assumes no interruptions in sleep from disturbances from the environment, such as sleep at home or a quiet and comfortable hotel **Good:** The environment that on average results in two interruptions per hour that each cost 5 min of sleep time, which is 50 min of restorative sleep per hour or 83% of Excellent. This is the level assumed for a Class I rest facility **Fair:** The environment that results in four interruptions per hour which is 40 min of restorative sleep per hour or 67% of Excellent. This is the level assumed for a Class II rest facility **Poor:** The environment that results in six interruptions per hour which is 30 min of restorative sleep per hour or 50% of Excellent. This is the level assumed for a Class III rest facility	Excellent

Appendix 3: Variables list

The following list contains all variables that were used in the final dataset. The first column contains abbreviations of the variables; the second column describes the variable names as used in the data analyses software programs; and the third column provides a full description of the variables.

Variable name	*Variable name used in analyses*	*Descriptions/Remarks*
General		
Crew_ID	CrewID	Crew member number
Duty_ID	DutyID	Duty number
WP_ID	WPNumber	Work period number. A work period is defined as the period between two recovery periods. A recover period is defined as a period of at least 36 hours containing at least two local nights (from 22:00h to 08:00h). Local means the time zone at the location the crew member is during the recovery period
FDP related variables		
StartH	TimeDayStartAT	Hour at which the duty starts in the time zone the crew member is acclimatised to
EndH	TimeDayEndAT	Hour at which the duty ends in the time zone the crew member is acclimatised to
FlightDur	FDPFlightDuration	Total flight minutes within FDP
FDPdur	FDPDuration	Total duty minutes in FDP
FlightDurWk FlightDur2Wk FlightDur4Wk	FlightTime7 FlightTime14 FlightTime28	Total flight minutes in 7, 14 and 28 preceding days
FDPWOCLm	MWOCL	Total time in WOCL (in minutes) in FDP (WOCL is between 02:00h and 05:59h)
FDPWOCL%	PWOCL	Percentage of the FDP that takes place during WOCL (between 02:00h and 05:59h). Thus, the following formula was used: MWOCL / FDPDuration
FPPaltWOCLm	FDPMCLow25	Total time in alternative WOCL (in minutes) in FDP (WOCL is now between 02:00h and 04:59h)
FDPaltWOCL1h	WOCL50	At least 1 hour of the FDP during alternative WOCL (WOCL is now between 02:00h and 04:59h)
FDPsectors	FDPSectors	Total number of sectors flown in current FDP
RestPeriod	RestPeriod	Total time off (in minutes) directly prior to FDP that is longer than 10 hours, but shorter than a recovery period
TimeBreak	TimeBSector	Total time (in minutes) between sectors within FDP
TimeOff	TimeOff	Total time (in minutes) in recovery period preceding work period
TZ_EW TZ_WE	TimeZonesEW TimeZonesWE	Number of time zones crossed from East to West (or vice versa) when comparing start with end of FDP
Km_SN Km_NS	DistSN DistNS	Kilometres from reference location to most Southern location of FDP or vice versa
HaulType	FDPHaulType	Haul type within FDP. Haul types are grouped as follows: 0. Cargo 1. Regional 2. Short-haul 3. Medium-haul 4. Long-haul If a duty contains a regional and short-haul flight, the haul type is 2
FDPmonth	FDPMonth	Month of start of FDP

Sleep-wake variables

SleepPred24 SleepPred48	HSleepPred24 HSleepPred48	Total predicted sleep (in minutes) in the 24 and 48 hours prior to TOD based on SAFTE model
SleepOpp24 SleepOpp48	HSleepOpp24 HSleepOpp48	Total sleep opportunity (in minutes) in the 24 and 48 hours prior to TOD based on BAM
SleepOppD24 SleepOppD48	HSleepOppD24 HSleepOppD48	Total sleep opportunity in darkness (in minutes) in the 24, 48 hours prior to TOD based on BAM
AwakeTOD	HAwake	Number of continuous minutes of wakefulness since the last period of sleep prior to TOD
AwakeStart	HAwakeM	Number of continuous minutes of wakefulness since the last period of sleep before start of the FDP
OutOfPhase	OutOfPhase	Sleepiness scale calculated with SAFTE model. SAFTE describes this scale as: *The difference between the environmentally driven circadian phase based on sleep pattern in the local time zone (Goal Phase) and the person's current physiological circadian phase (Current Phase). A positive value reflects a phase delay (westward travel) and a negative value reflects a phase advance (eastward travel). More than 3 hours is a fatigue factor: Sleep that occurs at local bedtime when the person is out of phase will be less restorative than when the person is not out of phase*

Fatigue measurement variables

EffectTOD	EffTOD	Effectiveness score calculated at TOD with SAFTE model. TOD was defined at half an hour before wheels on ground. SAFTE describes this scale as: *Effectiveness represents speed of performance on the Psychomotor Vigilance Test, scaled as a percent of a fully rested person's normal best performance (20 – 103). Effectiveness corresponds to the speed of cognitive performance, is highly sensitive to fatigue, and correlated with many other cognitive performance metrics. The higher the score the lower the fatigue risk* *Note: It is possible for scores to be greater than 100%. An example of such can occur as a result of a good night of rest and circadian function, followed by an afternoon nap*
EffectLow	EffMIN	Lowest level of effectiveness during FDP. For a detailed description of the effectiveness scale, see Effectiveness TOD above
CASTOD	CASTOD	Common alertness scale calculated at TOD with BAM. Jeppesen describes this scale as: *Scale with a range from 0 to 10,000 where 0 is the least alert state and thereby the highest fatigue risk. The CAS scale is directly anchored to KSS in a way CAS 0 = KSS 9, CAS 10,000 = KSS 1*
TimeLowEffect	FDPMinsEffLow	Total time (in minutes; at least 90 minutes) within the FDP during which the effectiveness score was below 87.8. This is equivalent to a KSS score of 5 or higher
TimeDayTOD	TimeofDayFatTOD	The time of day at the time the fatigue was calculated at TOD. This is the time the aircrew is acclimatised to at the time of measurement
TimeDayMAX	TimeOfDayFatMAX	The time of day at the time the highest fatigue was calculated. This is the time the aircrew is acclimatised to a the time of measurement
MinPassedTOD	FDPPassedTOD	Duty time passed (in minutes) at TOD

Grouping variables

FDP1LN	FDP2	At least one minute of the duty takes place between 02:00h and 04:59h and the duration of the duty is at least 10 hours. FDP1 (Night duties > 10h) was originally indicated as FDP2 in D1 Addendum; this FDP was ranked first, therefore FDP1 was introduced
FDP2DS	FDP6	One of the following variables is applies: FDP6ETES, FDP6ETLF, FDP6LTES, FDP6LTLF, FDP6Night. FDP Disruptive Schedules was originally indicated as FDP6 in D1 Addendum; this FDP was ranked second, therefore FDP2 was introduced

FDP2ETES	FDP6ETES	If the aircrew member works in a country or the airline's home base is situated in a country that is defined as late type and the FDP started between 05:00h and 05:59h
FDP2ETLF	FDP6ETLF	If the aircrew member works in a country or the airline's home base is situated in a country that is defined as late type and the FDP ended between 23:00h and 01:59h
FDP2LTES	FDP6LTES	If the aircrew works in a country or the airline's home base is situated in a country that is defined as late type and the FDP started between 05:00h and 06:59h
FDP2LTLF	FDP6LTLF	If the aircrew works in a country or the airline's home base is situated in a country that is defined as late type and the FDP ended between 00:00h and 01:59h
FDP2ES	FDPESFTL	If FDP6ETES or FDP6LTES
FDP2LF	FDPLFFTL	If FDP6ETLF or FDP6LTLF
FDP2N	FDP6Night	If at least one minute of the FDP took place between 02:00h and 04:59h
VeryES	FDPVES	Start FDP between 05:00h and 05:59h
ES	FDPES	Start FDP between 06:00h and 06:59h
LF	FDPLF	End FDP between 23:00h and 00:29h
VeryLF	FDPVLF	End FDP between 00:30h and 01:59h
ES2LT	ESLFFTL	Transition from early start (FDP6ETES or FDP6LTES) to late finish (FDP6ETLF or FDP6LTLF). These are consecutive duties
ES2N	ESNFTL	Transition from early start (FDP6ETES or FDP6LTES) to night (FDP6Night). These are consecutive duties
ES2LF	LFESFTL	Transition from late finish (FDP6ETLF or FDP6LTLF) to early start (FDP6ETES or FDP6LTES). These are consecutive duties
LF2N	LFNFTL	Transition from late finish (FDP6ETLF or FDP6LTLF) to night (FDP6Night). These are consecutive duties
N2ES	NESFTL	Transition from night (FDP6Night) to early start (FDP6ETES or FDP6LTES). These are consecutive duties
N2LF	NLFFTL	Transition from night (FDP6Night) to late finish (FDP6ETLF or FDP6LTLF). These are consecutive duties
ForwardR	WPRotCForw	Number of forward (or clockwise) rotations in consecutive FDPs. Forward rotations are: ESLFFTL, LFNFTL and NESFTL
BackwardR	WPRotCCBackw	Number of backward (or counter clockwise) rotations in consecutive FDPs. Backward rotations are: ESNFTL, LFESFTL and NLFFTL
NoConES	CONES	Number of consecutive early starts (FDP6ETES or FDP6LTES). If CONES = 2, this means that the previous FDP started early and that the current FDP started early.
NoConLF	CONLF	Number of consecutive late finishes (FDP6ETLF or FDP6LTLF). If CONLF = 2, this means that the previous FDP ended late and that the current FDP ended late
NoConN	CONN	Number of consecutive nights (FDP6Night). If CONN = 2, this means that the previous FDP was a night duty and that the current FDP is a night duty

Individual-related variables

Function	Function	Function of the crew member, this can be flight or cabin crew
RosterType	DataType	Type of roster that was analysed: worked or planned

European Commission