NASA/TM-2011-215972

First Annual Report

NASA-ONERA Collaboration on Human Factors in Aviation Accidents and Incidents

Ashok N. Srivastava,
Ames Research Center
Moffet Field, California

Patrick Fabiani
ONERA Aerospace Lab
Toulouse, France

July 2011

NASA STI Program... in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA scientific and technical information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NASA Aeronautics and Space Database and its public interface, the NASA Technical Report Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA Programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counter part of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collects papers from scientific and technical conferences, synopsia, seminars, or other meetings sponsored or co-sponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include creating custom thesauri, building customized databases, and organizing and publishing research results.

For more information about the NASA STI program, see the following:

- Access the NASA STI program home page at *http://www.sti.nasa.gov*

- E-mail your question via the Internet to help@sti.nasa.gov

- Fax your question to the NASA STI Help Desk at (301) 621-0134

- Phone the NASA STI Help Desk at (301) 621-0390

- Write to:
  NASA STI Help Desk
  NASA Center for AeroSpace Information
  7115 Standard Drive
  Hanover, MD 21076-1320

NASA/TM-2011-215972

First Annual Report

NASA-ONERA
Collaboration on Human Factors in Aviation Accidents and Incidents

Ashok N. Srivastava,
Ames Research Center
Moffet Field, California

Patrick Fabiani
ONERA Aerospace Lab
Toulouse, France

National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California, 94035-1000

July 2011

Available from:

NASA Center for Aerospace Information
7115 Standard Drive
Hanover, MD 21076-1320
443-757-5802

National Technical Information Service
5285 Port Royal Road
Springfield. VA 22161
703-487-4650

Ames Research Center

First Annual Report

NASA-ONERA Collaboration on Human Factors in Aviation Accidents and Incidents

Ashok N. Srivastava, PhD., Project Officer for NASA

Patrick Fabiani, PhD., Project Officer for ONERA

March 2011

Prepared by:

NASA Ames Research Center, Code TH, Mountain View, CA

ONERA, Toulouse, France

Stinger Ghaffarian Technologies Inc., Mountain View, CA

Executive Summary

This is the first annual report jointly prepared by NASA and ONERA on the work performed under the agreement to collaborate on a study of the human factors entailed in aviation accidents and incidents, particularly focused on the consequences of decreases in human performance associated with fatigue. The objective of this agreement is to generate reliable, automated procedures that improve understanding of the levels and characteristics of flight-crew fatigue factors whose confluence will likely result in unacceptable crew performance. This study entails the analyses of numerical and textual data collected during operational flights. NASA and ONERA are collaborating on the development and assessment of automated capabilities for extracting operationally significant information from very large, diverse (textual and numerical) databases; much larger than can be handled practically by human experts.

This report presents the approach that is currently expected to be used in processing and analyzing the data for identifying decrements in aircraft performance and examining their relationships to decrements in crewmember performance due to fatigue. The decisions on the approach were based on samples of both the numerical and textual data that will be collected during the four studies planned under the Human Factors Monitoring Program (HFMP). Results of preliminary analyses of these sample data are presented in this report.

Most of the milestones scheduled for completion during this first year were achieved in essence, if not precisely in the expected detail. Limitations in the accomplishments were due, in part, to delays at easyJet in developing the software needed to put the raw flight-recorded data into a format compatible with NASA's algorithms for analyses That issue was resolved and NASA has been accessing easyJet's flight data on a daily basis since April 1, 2010. Another factor that limited some of the accomplishments was that ONERA and NASA were not working with precisely the same set of data from easyJet. This situation was brought to easyJet's attention and will be rectified prior to the next flight-crew study, which is currently planned for 2011.

1. Introduction

Background

Human fatigue has been recognized as a significant factor in aviation accidents and incidents. Maximizing alertness and performance levels during aviation operations is critical to maintaining the continued safety of the air transportation system. Flight-crew fatigue, especially in short-haul (regional) operations, has been recognized worldwide as a frequent contributing factor in incidents and accidents. There is a need to develop scientifically valid fatigue-management approaches that lead to continuous safety enhancements by identifying and addressing both physiological and operational fatigue factors across time and changing circumstances.

In recognition of the important influence of fatigue on crew performance in commercial airline operations, NASA entered into collaboration with easyJet Airline Company Ltd. on the Human Factors Monitoring Program (HFMP) study. (Reference the HFMP document and the First Annual easyJet Report TM.) The HFMP includes a series of studies being conducted to better understand how both latent and proximate causal fatigue factors potentially contribute to impaired flight and cabin-crew performance. NASA is collaborating with easyJet on the HFMP studies by providing technologies and methodologies to enable a data-driven and scientifically-based process that supports easyJet's Fatigue Risk Management System (FRMS).

Since the mid-1990's, NASA and the Office National d'Études et de Recherches Aérospatiales (ONERA) have collaborated on studies to understand human factors in aeronautical operations and incidents, which have been viewed as very beneficial by both agencies. NASA's agreement with easyJet offered a new opportunity to continue this excellent collaboration and to expand the research into a new, but directly related, aspect.

ONERA has entered into a similar agreement with easyJet so that this important research to be conducted under the HFMP will benefit from the highly successful collaboration between NASA and ONERA on developing and validating methodologies with which to extract and fuse information from large, diverse data sources to assist aviation safety analysts in:

1. Discovering expected and unexpected events or trends in system performance that could compromise the safety of the system
2. Identifying latent and proximate causal factors of those events to enable data-driven decisions on interventions or mitigations

The process to be jointly developed by NASA, ONERA, and easyJet will allow easyJet to monitor and manage safety risks associated with fatigue-related error. This process could serve as the foundation of FRMS within other air-carrier operations.

Project Overview

A goal of this collaboration is to understand (or, at least, to set up procedures with which to understand) the levels and characteristics of those latent and proximate factors of flight and cabin-crew fatigue whose confluence will more likely result in unacceptable crew performance and, in turn, anomalous aircraft operation.
The three-way collaboration is represented in Figure 1.1, which not only indicates the shared responsibilities, but also reflects the three separate agreements between NASA and ONERA, NASA and easyJet, and ONERA and easyJet for the sharing of data and informat

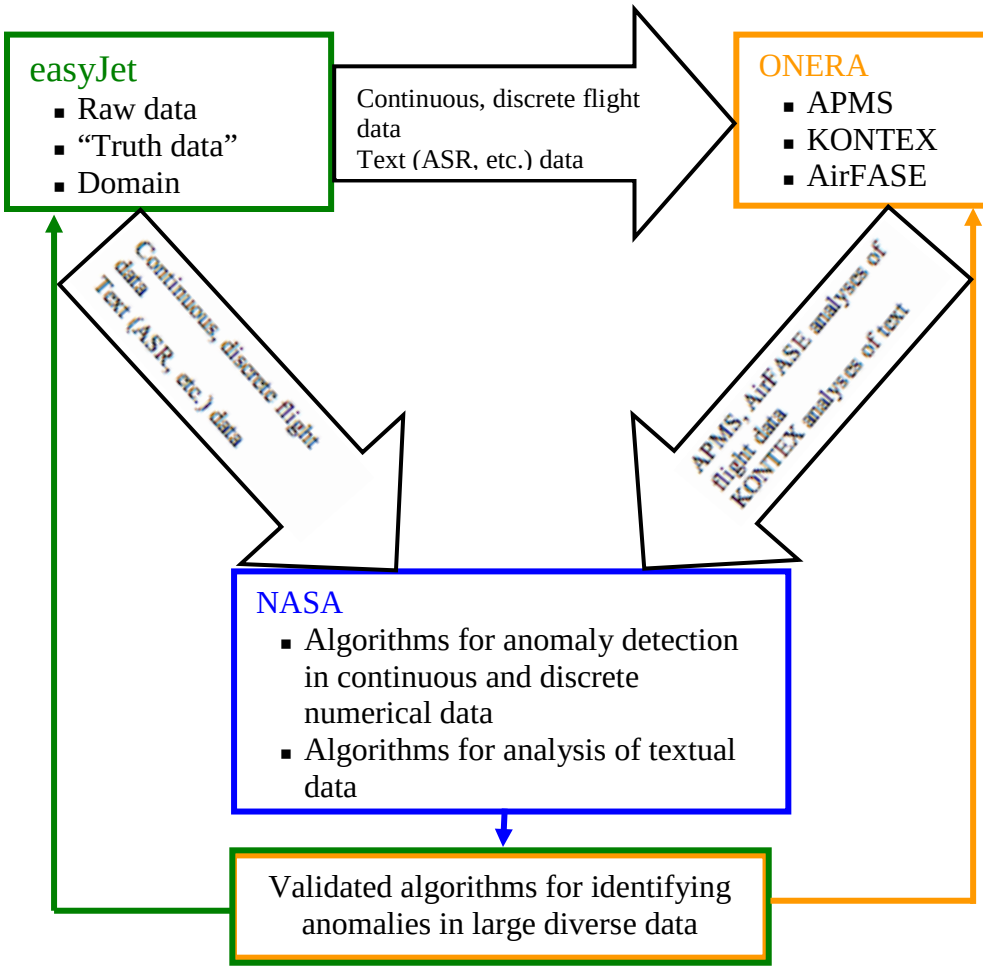

Figure 1.1: The ONERA-NASA-easyJet

Collaboration

The work being performed jointly by NASA and ONERA under this agreement relies on the acquisition of data to be provided by easyJet Airlines Company Limited to both NASA and ONERA under separate agreements. EasyJet's plan for a series of four experiments is described in detail in their publication, Human Factors Monitoring Program (HFMP), Fatigue Risk Management Scientific Study, Methodology; December 3, 2009.

Coincidentally, ONERA has become a part of a small consortium established by the French Government to gain understanding of the relationship between flight crew schedules in short-haul (regional) operations and flight crew fatigue, and the consequent impact on flight crew performance. ONERA has access to linked textual and numerical data that could provide information related to flight crew fatigue. However, ONERA's analytical capabilities for extracting information from numerical and textual data are limited and highly labor intensive, while NASA has been developing automated analysis tools to extract information from large sets of numerical and textual data. The results of the ONERA analyses will be used to help in the development and validation of NASA's automated tools.

The HFMP entails acquiring, processing, integrating, and interpreting large quantities of diverse numerical and textual data collected from flight-crew and cabin-crew participants during easyJet's normal operations with different scheduling strategies. Under the terms of a Space Act Agreement (SAA) with easyJet, NASA has been granted access to the following data: 1) aircraft

performance, 2) individual demographic information, 3) subjective questionnaires, 4) individual physiological measures, and 5) objective neurocognitive variables. NASA is considering these data from two perspectives.

- Personnel of the Intelligent Systems Division (Code TI) will analyze aircraft-performance data (i.e., flight-recorded data and Air Safety Reports)
- Personnel of the Human Systems Integration Division (Code TH) will analyze individual crew-performance data (i.e., demographic information, questionnaires, and physiological and neurocognitive data)

The results of the two studies would also be combined to relate decrements in aircraft performance to decrements in crew performance.

Code TI's role is to develop and apply algorithms that automatically extract information on anomalous events from the flight-recorded data and to fuse that information with information extracted automatically from the easyJet's Air Safety Reports (ASR). The developed algorithms will contribute to NASA and easyJet's joint efforts to develop efficient and reliable methodologies that can extract and merge information from large, diverse data sources to assist aviation safety analysts with: 1) identifying expected and unexpected events or trends in system performance that could compromise the safety of the system; and 2) identifying the latent and proximate causal factors of the events identified to enable data-driven decisions on interventions and mitigations.

ONERA's role is primarily with complementary analyses of aircraft performance data in collaboration with NASA personnel in Code TI.

At the same time, Code TH's role is to analyze physiological and neurocognitive measures, collected as part of the easyJet data-collection process to track individual levels of fatigue and performance across the entire time-course of the duty cycle. Individual fatigue and performance levels will be further analyzed in conjunction with personal profile variables to identify potential causal factors of fatigue. ONERA is not expected to participate in these analyses, but will join in the integration and assessment of the results of all the analyses to support common goals of this research with easyJet.

EasyJet is seeking answers to the following specific questions:

1. Is there reliable evidence that levels of fatigue can be correlated with scheduling strategy *AND* is there reliable evidence that scheduling strategy is a causal factor of performance-degrading levels of fatigue?
2. What is the minimum set of measures to reliably indicate that identified aircraft performance decrements were probably related to fatigue?
3. Pragmatically, which measures can be implemented during normal operations to monitor for levels of human fatigue that could affect performance?
4. What are the data sources that provide reliable information on the consequences of performance-degrading levels of fatigue?
5. What are the data sources that provide reliable information on the latent and proximate causal and contributing factors of human fatigue?
6. What are the fatigue profiles of operators based on individual measures over the course of a flight? Are these indicators convergent?

The overall goal of this research is to identify the simplest reliable measurement system for monitoring fatigue, crew performance, and aircraft performance. This will be accomplished

by minimizing the number and complexity of a combination of objective measurements that are shown to have reliable associations with changes in performance and fatigue levels.

At NASA, this work has become incorporated within the Data Mining and Knowledge Discovery Theme and the Human Systems Solutions Theme of NASA's System-wide Safety and Assurance Technologies (SSAT) Program.

At ONERA, this work is being perfomed within the Département Commande des Systèmes et Dynamique du vol (System Control and Flight Dynamics Department: DCSD).

This report focuses on work related to analyzing flight-recorded data and Aviation Safety Reports (ASR) because these are the areas of primary collaboration between ONERA-DCSD and NASA Ames Code TI. The work being performed by Code TH on the physiological and neurocognitive measures is reported in annual reports prepared jointly by NASA and easyJet. (See, for example, the NASA TM reference to the first annual easyJet report).

During the first year of the ONERA-NASA agreement, the Project Officers met twice. This is primarily a report of the status of work as presented during the 2nd meeting that was held at NASA Ames Research Center in Mountain View, California, July 21-23, 2010.

The reports presented at this meeting were largely preliminary because:

(1) Not all of the data collected during the first experiment with flight crews from September 2, 2009 to October 3, 2009 had been received by both NASA and ONERA at that time.

(2) There were known deficiencies in the data acquired during the first experiment.

Therefore, the available data were used by both NASA and ONERA largely to assess the quality and quantity of the data to be expected in the next three experiments; to identify any additional deficiencies that could limit or prevent the achievement of the objectives of the NASA, ONERA, and easyJet agreements if not resolved; and to develop the approach to be pursued by NASA and ONERA.

There is very little more to report since that meeting because no data needed for the objectives that are relevant to aircraft performance are expected until the next flight-crew study currently scheduled to commence in April 2011.

Data Collection

The HFMP calls for data collection on 22 subjects during each of the four studies conducted by easyJet. Flight crews are the subjects for studies 1, 3, and 4, and cabin crews for study 2. Each subject is to be either on Flexible Roster Variation (FRV) or the Fixed-Pattern Design (FPD) schedule, as diagramed in Figure 1.2.

				Block A								
	D/O	D/O	D/O	E1	E2	E3	L1	L2	D/O	D/O	D/O	
Block B							**Block C**					
E1	E2	E3	L1	L2	D/O	D/O	E1	E2	E3	L1	L2	

FRV Schedule of 23 Consecutive Duty and Off Days

D/O=Day Off; E=Early Departure; L=Late Arrival

				Block A							
D/O	D/O	D/O	D/O	E1	E2	E3	E4	E5	D/O	D/O	D/O
Block B					**Block C**						
L1	L2	L3	L1	L2	D/O	D/O	D/O	D/O			

FPD Schedule of 21 Consecutive Duty and Off Days

Figure 1.2: Schedules of Duty Days and Off Days

The following data are to be collected during the four studies:

Aircraft Performance Data

- In-flight recorded data – trace records and exceedances
- Air Safety Reports (ASR's)

Individual Crew Demographic Information

- Schedule data (rosters)
- Demographic data – age, commute distance, base, flying experience, family status, etc.

Individual Crew Physiological and Cognitive Data

- Samn-Perelli – subjective alertness scale
- Mood scale – subjective
- Fatigue countermeasures employed per sector
- PVT – psychomotor vigilance task
- Sleep data-actigraphy (Actiwatch)
- Sleep diary information
- NASA TLX questionnaire data
- Hassle factors questionnaire data
- Morningness/Eveningness Questionnaire (MEQ)
- Checklist of Individual Strength (CIS)
- Epworth Sleepiness Scale (ESS)
- Bio-harness data (Electrocardiography: ECG; Respiration Rate: RR; Skin Temperature: ST; Posture: indicates upright or supine body position; and Physical Activity: three axis accelerometer)
- Crew portal questions on sleep quantity and alertness pre/post duty

All data are annotated with a common timestamp (GMT) to enable their linkage. These measures are fully explained in Stewart (2009). The in-flight-recorded data are equivalent to data

collected under Flight Operational Quality Assurance (FOQA), which is the name given to the program in the U.S. that entails continuously recording, monitoring, and analyzing data on hundreds of flight parameters recorded during flight. Recording for FOQA is performed in parallel with the system associated with the mandated "black box" from which data are typically analyzed after accidents.

2. Aircraft Performance

Background

NASA's Code TI and ONERA's DCSD will collaborate on developing, validating, and applying algorithms that automatically extract information on anomalous events from the flight-recorded data and from easyJet's Air Safety Reports (ASR). The aim is to identify relationships between anomalous events and/or exceedances identified in the flight data and their possible causes identified in the ASR's. It is recognized that not all decrements in aircraft performance are caused by flight-crew fatigue, but that the occurrence frequency of some aircraft–performance markers can be causally correlated to the fatigue level of the flight crew.

NASA and ONERA will try to identify events in flight-recorded data that are reliably causally correlated to a level of flight-crew fatigue and could serve as surrogates for indications of fatigue in physiological data collected during some flights. The idea is to use these surrogates in the clear majority of flights for which physiological data cannot be collected. The developed algorithms will contribute to ONERA and NASA's joint efforts to develop efficient and reliable methodologies to extract and merge information from large, diverse data sources. The goals of this research are to assist aviation safety analysts to:

1. Identify expected and unexpected operationally significant events or trends in aircraft performance that could compromise the safety of the system.
2. Identify the latent and proximate causal and contributing factors of the events identified to enable data-driven decisions on interventions or mitigations.
3. Identify events where flight-crew fatigue was a contributing or causal factor.

Approach to Flight-Data Analysis

The FOQA programs that have been implemented by most air carriers worldwide represent the state-of-the-art for analyzing flight-recorded data. Software venders have developed methods for automatically searching these data for expected events, i.e., an event that has been predefined as, typically, a single variable that has exceeded a value during a particular phase of flight, and one that is considered outside of the established Standard Operating Procedures (SOP's).

The purpose of NASA Code TI's work in flight-data analyses is to discover the unexpected events that could compromise the safety of operations to complement and supplement the search for the expected events performed under the FOQA programs.

The algorithms developed by Code TI search for sets of continuous parameters and binary switches that contribute to an event that is considered statistically anomalous in a multivariate comparison with normal operations.[1] The automatic identification of the contributions of the particular continuous and discrete parameters entailed in the identified anomalous event assist the domain expert in ascertaining its operational significance. The

[1] Safe "normal" operations are not always completely consistent with SOP's.

algorithms are designed to process very large data sets (collected at the rate of over 10,000 Kb per flight and about 2.5 Tb per year) in nearly real time.

An outline of the approach, which has been named the Multiple Kernel Anomaly Detection (MKAD), for detecting the unexpected anomalous events in the numerical flight-recorded data is diagrammed in Figure 2.1. This is a description of the approach that we would follow using algorithms that have already been developed. However, work on improvements is ongoing and the most recent algorithms available at the time of final analyses will be used.

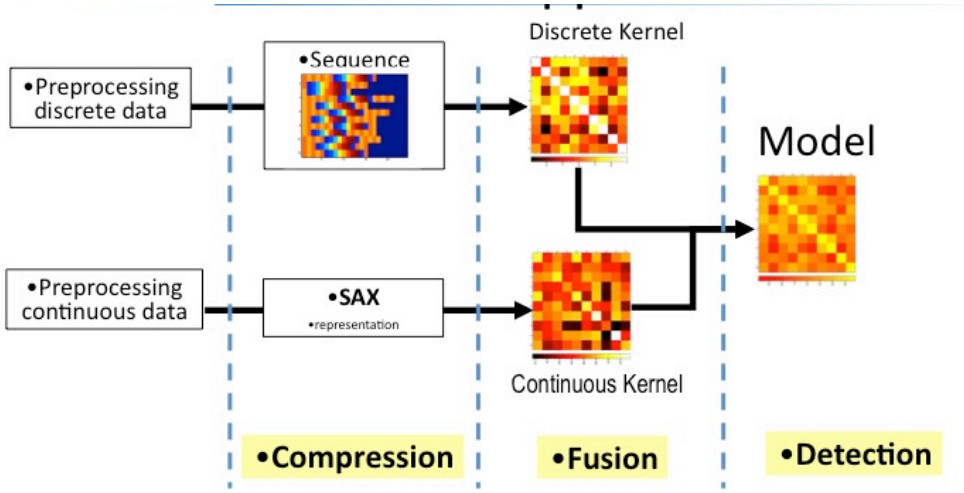

Figure 2.1: Outline of Approach

Initially, discrete and the continuous data are processed independently. Processing of discrete data is achieved using processes based on the previously developed algorithm called sequenceMiner, which detects and characterizes anomalies in large sets of high-dimensional symbol sequences that arise from recordings of switch sensors in the cockpits of commercial airliners. SequenceMiner works by performing unsupervised clustering (grouping) of sequences using the normalized longest common subsequence (LCS) as a similarity measure, followed by a detailed analysis to detect anomalies. An outlier sequence is defined as a sequence that is far away from the cluster center. The algorithms provide a coherent description to an analyst of the anomalies in the sequence when compared to more normal sequences.

Processing of the continuous data uses **S**ymbolic **A**ggregate appro**X**imation (SAX - invented by Eamonn Keogh and Jessica Lin in 2002), which provides a symbolic representation for time series that allows for dimensionality reduction and indexing with a lower-bounding distance measure.

These steps are followed by a fusion process in which discrete and continuous kernel functions are developed. These kernel functions provide a measure of how similar a new, previously-unobserved flight's data is to most (presumably normal) previously-observed flight data. The discrete and continuous kernels are fused into a single model that is used for detection of anomalous events associated with multivariate (continuous and discrete) parameters. These anomalous events are then compared with expected events found by the search for prescribed exceedances.

Currently, the FOQA-like exceedances are identified for easyJet by Airbus and its contractor using its proprietary software: AirFASE. Through existing agreements among Airbus, easyJet, and ONERA, ONERA will undertake the task of identifying exceedances in the flight

data that are collected during HFMP experiments. Through existing agreements among ONERA, easyJet, and NASA, these exceedances will be made known to NASA. NASA Code TI and ONERA personnel, with assistance from subject-matter experts, will compare the results of the search for the unexpected events using NASA's anomaly detection algorithms with the exceedances identified by ONERA using AirFASE. Based on previous similar experiments, the comparisons are expected to fall into three categories. In the first category, there will be anomalous events that are identical to those prescribed exceedance events found using AirFASE. The multivariate information obtained using the algorithms for anomaly detection will complement the single variable exceedance information and give the safety analyst a better understanding of the event.

There will be a second category of events that are identified as exceedances for which no anomalous events have been identified. Each of these is likely to have a different explanation. Reasons found in previous studies have had to do with exceedances based on computed parameters within the AirFASE processing that were not recorded and available to the search for anomalies, or exceedances that occur so frequently that they are not identified as anomalous events when compared to normal operations.

The third category will be events that have been identified as anomalous but have not been found using AirFASE. Figure 2.2 indicates the types of events that might be discovered based on searching the combination of continuous and discrete variables.

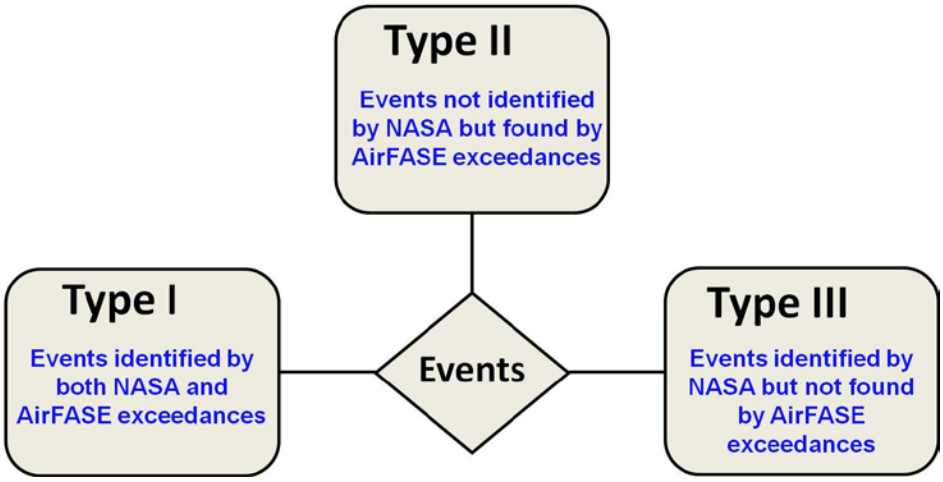

Figure 2.2: Types of Events

Events in all of these categories, along with the identification of the parameters that caused them to be considered statistically significant events, will be of value to subject matter experts in identifying events considered to be operationally significant. Newly discovered events that are deemed operationally significant could then be used to define a pattern to be used in a routine search of past or future flights for prescribed events.

During the first year of this collaboration, ONERA undertook preliminary work on extracting easyJet's FOQA exceedances from continuous flight data using AirFASE, on identifying anomalies using APMS's Morning Report of Atypical Flights, on trying to use sequenceMiner on discrete data, and on plans for comparing the results of these studies with analyses by NASA using new algorithms for detecting anomalies.

The work with sequenceMiner or an equivalent algorithm for searching for anomalies in discrete data is of particular interest to the objectives of this research. Previous studies have

indicated that sequenceMiner identified anomalous sequences in the operations of switches on the flight-management/autopilot system during the approach phase that domain experts had said were indications of mode confusion. ONERA and NASA will jointly select a subset of easyJet's flight data, a set of parameters, and the phase(s) of flight to be used in their respective analyses to search for indications among discrete variables of mode confusion. ONERA will also analyze the selected data set using their previously developed manual techniques to search for indications of mode confusion; then NASA and ONERA will collaborate on analysis to establish whether mode confusion, as indicated by anomalies in certain discrete variables using sequenceMiner or its equivalent, can be reliably and causally correlated with other evidence of flight-crew fatigue developed in this research.

During the course of discussions at the second meeting of the Project Officers, it was found that ONERA and NASA were not working with precisely the same set of data from easyJet. This situation was brought to easyJet's attention and will be rectified prior to the next flight-crew study. Also, ONERA and NASA had not yet agreed on the definitions of the phases of flight to be used by both in their analyses. This too is being resolved.

Approach to Air Safety Report Analysis

Code TI's work on text analysis is to develop algorithms to extract information from the textual data of easyJet's Air Safety Reports, which indicate the flight-crew's perspective on factors entailed in the reported event. In particular, the search of the ASR's will be to find reports where the crew has identified human fatigue as a contributing or causal factor. NASA's approach is based on extracting Topics from free-form text using Latent Dirichlet Allocation (LDA). Frequently, more than one Topic is associated with each event. A Topic is defined by a set of words. The number of Topics to be identified in a set of textual reports is pre-specified, as are the number of words to define a Topic.

The assumptions of the LDA are that:

A document (D) is a mixture of multiple topics (T)

– A topic (T) has several words associated with it

– A document is generated by:

1. First selecting a topic
2. Selecting a word associated with that topic
3. Repeating steps 1 and 2 for as many words as desired to form that document

– Collection of such documents forms a corpus (C)

The problem for LDA is, given C (a data set of textual reports), to find T (a set of topics representative of C). NASA's process was tested on a set of 66,311 ASRS reports composed of 59 defined events (ASR anomaly categories), and the following are three of the Topics that were found:

TOPIC 1	vert	pwr
autoplt	ctl	
acft	disconnected	**TOPIC 2**
spd	selected	time
capture	fpm	day
mode	light	leg
rate	clb	contributing
level	pitch	factors
engaged	manually	hrs
leveloff	warning	crew

factor	alerter	missed
fatigue		clred
night	**TOPIC 3**	msl
trip	apch	intercept
rest	rwy	vectored
duty	visual	sight
flying	ils	gar
long	twr	terrain
late	lndg	field
previous	loc	uneventful
incident	arpt	ctl
lack	final	

Each of these Topics reflects fatigue as being, at least, a contributing factor. Examples of ASRS anomalous events in which fatigue-related Topics such as these were involved were:

- Altitude Deviation
- Spatial Deviation
- Ramp Excursion
- Landing with Clearance
- Runway Incursion
- Unstabilized Approach

In the absence of a well-defined categorization of easyJet's ASR's into prescribed anomaly classes, the rich categorization of ASRS reports will be leveraged and the knowledge learned from these reports will be transferred to categorizing ASR's into the ASRS anomaly categories. The expertise that has already identified certain ASRS anomalous events with fatigue will be the basis of determining which of the Topics identified by LDA are fatigue-related with a high probability, and also which of the Topics occurred in a particular ASR event. The approach would be to develop a classifier (e.g., Mariana) and a clustering technique (e.g., LDA) that have learned from the ASRS database, and then test all of the ASR documents with both the classifier and clustering technique to find documents where (1) fatigue is a high probability, and that address LDA topics related to fatigue; and (2) fatigue is a high probability, and that relate to various ASRS anomaly categories that have been identified with fatigue by ASRS analysts. The point of this is to find reports that indicate both fatigue and some safety problem. This would allow the transfer of the information learned from ASRS to ASR and the verification, if similar documents exist, in both. Of course, subject-matter experts at easyJet will provide final confirmation and validation of the findings.

ONERA's role in this collaboration is primarily concerned with the aircraft performance based on the analyization of the numerical continuous and discrete flight-recorded data and textual Aircraft Safety Reports (ASR's) generated by easyJet. However, during the meeting in July 2010, ONERA also reported on a relevant study that had been performed of human fatigue issues for French short-haul transport aircraft in support of an effort by the French government and airline industry to define a fatigue risk management system similar in concept to that of easyJet. ONERA used a simple bio-mathematical model of fatigue that only considered crew working hours for each flight to predict the risk of fatigue associated with each flight. ONERA then examined whether this level of fatigue correlated statistically with the occurrence of FOQA exceedances. The conclusions from this study were:

- Some FOQA events are sensitive to the fatigue level of the crew, but not all, and not all in the same way. So it is necessary:

    - To identify these events
    - To characterize how fatigue influences their occurrence
- A study of several possible factors shows that the variable produced by the bio-mathematical model is the most discriminating
- Selected events could be used as fatigue markers in a fatigue risk management system
- More sensitive events could be designed

ONERA considered these to be highly tentative conclusions because there are likely significant individual differences that this model did not take into account. The physiological and neurocognitive measures obtained during the easyJet studies to be analyzed by NASA Code TH personnel will enable an evaluation of ONERA's study and a comparison with results from other models.

Approach to Fusing Information From Flight Data and ASR's

NASA and ONERA will collaborate on exploring the relationships between the information extracted from the flight-recorded data and information from the ASR's. While flight data tells us *what* happened, the ASR's might give us clues as to *why*. Our focus for this study will be on indications in ASR's that fatigue might have been a factor in corresponding anomaly or aircraft performance events in numeric flight data. Of course, not all aircraft performance events are crew fatigue-related. Therefore, we will start with ASR's in which our analysis of the narrative has found a high probability that the crewmember believed fatigue was a factor. Using the flight identification and timestamp, we will relate information from each fatigue-related ASR to the event(s) that have been identified in the corresponding flight. Domain experts from easyJet will be asked to review our identified correspondences to determine if the statistical anomalies that we identify correspond to operationally significant anomalies. This use of ASR's to identify the potential of fatigue as having been a factor in an event will provide useful information for the analysis that will merge information on decrements of individual crew performance with decrements in aircraft performance.

The fusion method just described is largely a manual fusion method, in that the only automated parts are anomaly detections within the numeric data, and anomaly identifications within the ASR's. We are performing research to fully automate the process of finding correspondences between anomalies that appear in the numeric data and anomalies described in the text. NASA plans to extend the MKAD method described above to allow text reports for the same flight to be incorporated as available numeric data; thereby allowing MKAD to find flights that have anomalies in both numeric and text data, and to identify correspondences between anomalies in both types of data. As mentioned before, one difficulty to overcome is that most of the flight data do not have corresponding ASR's. We must ensure that the method that we use utilizes ASR's when available, but does not skew results when an ASR's are not available.

In support of the effort to develop methods for reliably fusing information extracted from numerical and textual data, ONERA will use manual analysis techniques they have developed, such as KONTEX, on subsets of linked textual and numerical continuous and discrete data for consequences and causal factors of flight-crew fatigue.

Under its separate agreement with easyJet, ONERA will gain access to subsets of the same datasets used for this development by NASA. The results of ONERA's analyses will be used for comparison with, and guidance for, NASA's automated analyses of large data sets of these linked sources of data.

Progress to Date on Analyzing Flight Data and ASR's

Tasks for this year have focused on the relatively small sample flight-recorded datasets received by NASA and ONERA to date and the resolution of issues of data transmission, data quality, and transformations needed to make the data format compatible with the analysis method. The work associated with extracting information from continuous and discrete numerical data and the work associated with extracting information from textual data have, so far, been performed independently.

Flight-recorded Numerical Data

As of the writing of this report, NASA has not yet received flight-recorded data collected during the first experiment with flight crews during September 2009. However, beginning on April 1, 2010, easy Jet made flight data available to NASA for data analysis on a daily basis. Since then, NASA has been able to collect over 209,000 flights (2.4TB) over 9 months and performed anomaly detection using algorithms that were developed at NASA. We used 47,000 flights that were available at the time of this study (mid-July) to test the MKAD methodology for anomaly detection. For purposes of this demonstration, the algorithm focused only on the landing phase of flight for analysis from 10,000 ft. to landing. The results are presented in Appendix A and are divided into three categories: Events that were found to be *Statistically Significant*, events that we have called *Heterogeneous*, and a large number of *Go-arounds*. The contributing parameters that were identified by the algorithm are listed along with a plot for each anomaly. No attempt has been made yet to relate these to FOQA exceedances or to assess their operational significance. Consideration is being given to a further study of the identified Go-arounds to see if it is possible to define the circumstances (precursors) that required a Go-around. Of course the information will not be available from the flight data if the Go-around was in response to a command from ATC.

Aviation Safety Report Analysis

LDA was applied to a set of 12,665 easyJet ASR's to identify Topics in the free text. Although many Topics were identified using the LDA, we focused our attention on a preliminary study on one Topic that might relate to crew fatigue. The words associated with this Topic are shown below:

TOPIC
duty
food
time
safety
sector
rest
night
delayed
hours

water
early
long
room
days
risk
sandwiches
period
hotel
days

The words in this Topic point in several interesting directions. One is the issue with duty hours being long. Then, some of the reports point to insufficient rest at night, either at home or in a hotel room. There is also a class of reports that talk about problems with the quantity and quality of crew food. In Appendix B, we present ten ASR's with the highest probability of being associated with this Topic.

As of the writing of this report, we have still not received the flight data from the first experiment. Therefore, we made no study of the linkage of the numerical data and the ASR reports. When we receive linked flight data and ASR's, we hope to match ASR's that we have identified with fatigue with the aircraft performance data to establish a causal relationship between any such report and the associated flight performance. The additional problem that we found in the ASR's that we received so far is that, as exemplified in the ten reports in Appendix B, the fatigue-related reports do not, for the most part, address any specific event that occurred (possibly as a consequence of fatigue), but rather speak of company policies that may have been contributing factors to fatigue.

3. Summary and Conclusions

The closing discussions on the final day of the meeting of the Project Officers at NASA Ames in July 2010 addressed the work the two organizations had done, and plan to do, and how the two agencies will continue to work together in the future to achieve the goals of the current agreement. Much time was spent discussing the many deficiencies in the data from the first experiment and how they might be resolved. The following specific action items were identified:

1. Proceed with formal agreements to ensure that ONERA receives the same data in the same format as NASA from the first completed experiment and the next three experiments. NASA's current agreement with easyJet says, *"The Parties will not share original raw data received from easyJet, but rather will share only the results of analyses of such data. If it becomes necessary to transfer easyJet data from one Party to another, such transfer will be made through easyJet."* Consequently, easyJet must be consulted on the preferred procedure to resolve this problem.
2. ONERA and NASA will decide on phases of flight to be used and their definitions.
3. NASA will provide ONERA with additional training on sequenceMiner.
4. ONERA and NASA will follow-up on easyJet's responses to resolve the issues

with the data that had previously been submitted.

5. ONERA and NASA will compile and submit to easyJet the needs for additional information that were identified during this meeting. These include the following:

    a. Statistical information (e.g., age, flying experience, gender, commute times) on the general populations of flight and cabin crews at easyJet to assess how well sample sets of subjects represent the full population.

    b. Information on whether cabin crews will submit ASR's in the next experiment.

    c. Definitions of the FOQA exceedances used in AirFASE.

The Project Officers agreed that, in view of the issues that both NASA and ONERA needed to resolve with easyJet, it would probably be highly worthwhile if the two of them went together to visit easyJet for a discussion of these problems.

4. Meeting with easyJet

A meeting of the NASA and ONERA Project Officers with easyJet was arranged for October 19, 2010. However, other commitments intervened and neither Dr. Srivastava nor Dr. Fabiani was able to participate in person, although both joined in a telecon with easyJet personnel in the afternoon. On the morning of that day, Dr. Michael Feary of NASA Code TH represented both Project Officers in a personal meeting with Messers Jim Pegram, Head of easyJet SMS; Phil Barton, Head of easyJet FRMS; Simon Henchie, Head of easyJet Flight Data and Rostering; Ms. Lydia Harbour, and several other data analysts.

Dr. Feary reported that easyJet representatives assured him of their commitment to perform the HFMP, but that there would be a delay until easyJet and their employees union had completed negotiations on a new contract. It was agreed that the unions had to be involved. Another issue that we thought had been resolved concerned the use of the Bio-harness that had been raised by easyJet's Occupational Health and Safety office, who were thinking of mandating a full medical profile of each subject before and after the study. This also needed to be resolved internally before easyJet could organize plans for the remaining studies. NASA's interest in having the objective measures obtained with the Bio-harness and the Actiwatch was understood and appreciated by easyJet.

NASA had already reported to easyJet that the first study of flight crews had yielded enough information on nine subjects to exercise the FAST model. The Imperial College, who had been engaged by easyJet to use the SAFE and FAID models, had applied these to the data from the same nine subjects and expressed high interest in comparing the results from the three models.

When Doctors Srivastava and Fabiani joined by telecon that afternoon, these same subjects were discussed and easyJet's commitment to completing the HFMP was reiterated. The NASA and ONERA Project Officers expressed their understanding of easyJet's current problems and offered whatever assistance they could provide in resolving them. They raised the issues regarding deficiencies in the data from the first study that had been identified during the meeting at Ames between NASA and ONERA. They received assurance that these would be corrected before the next experiment with

flight crews.

In telephone and e-mail communications subsequent to the meeting at easyJet, NASA was told that easyJet's upper management had given very strong support and encouragement to proceed with the HFMP study. NASA was assured that easyJet would give high priority to correcting all the deficiencies in the data and would ensure the use of the Bio-harness in the next studies. The easyJet Project Officer for the NASA-easyJet agreement sent an e-mail to the NASA Project Officer assuring him of easyJet's commitments.

NASA reported to easyJet that this study had elicited very high interest and support at a formal review of the System-wide Safety and Assurance Technologies (SSAT) project, of which this work was a part. NASA's invited presentation of this joint research was also well received by the panel of the National Academies of Engineering and Sciences as it related to the study they had undertaken of flight-crew fatigue.

NASA's Project Officer views the responses to these reviews as strong endorsements of this work and is encouraged that we, together, have in fact formulated a study that is important to the aviation industry.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
10-10-2020	Technical Memorandum	10-2009 - 9-2010

4. TITLE AND SUBTITLE

NASA-ONERA
Collaboration on Human Factors in Aviation Accidents and Incidents

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Ashok N. Srivastava and Patrick Fabiani

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER
534723.02.01.01.10

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

NASA Ames Research Center
Moffet Field, California 94035-1000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

National Aeronautics and Space Administration
Washington, DC 20546-0001

10. SPONSORING/MONITOR'S ACRONYM(S)
NASA

11. SPONSORING/MONITORING REPORT NUMBER
NASA/TM-2011-215972

12. DISTRIBUTION/AVAILABILITY STATEMENT

Unclassified-Unlimited
Subject Category 03
Availability: NASA CASI (443) 757-5802

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This is the first annual report jointly prepared by NASA and ONERA on the work performed under the agreement to collaborate on a study of the human factors entailed in aviation accidents and incidents, particularly focused on the consequences of decreases in human performance associated with fatigue. The objective of this agreement is to generate reliable, automated procedures that improve understanding of the levels and characteristics of flight-crew fatigue factors whose confluence will likely result in unacceptable crew performance. This study entails the analyses of numerical and textual data collected during operational flights. NASA and ONERA are collaborating on the development and assessment of automated capabilities for extracting operationally significant information from very large, diverse (textual and numerical) databases; much larger than can be handled practically by human experts.

15. SUBJECT TERMS

human factors, aviation safety, fatigue studies, data mining

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19b. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			STI Help Desk (email: help@sti.nasa.gov)
U	U	U	UU	22	19b. TELEPHONE NUMBER *(Include area code)* (443) 757-5802