

In dieser Ausgabe:

- Die Zukunft des Flugdatenaustausches: Flight Object Interoperability (FO IOP)
- Improving Climb Performance Prediction in Air Traffic Control with Machine Learning and Full Flight Simulator Verification
- Tactical inputs for a stripless ATC system via Multi-Touch Gestures
- Erkenntnisse aus dem Projekt "StayCentered - Methodenbasis eines Assistenzsystems für Centerlotsen"

-

Inhaltsverzeichnis

Übersicht der Artikel in diesem Heft	4
Editorial.....	5
Die Zukunft des Flugdatenaustausches: Flight Object Interoperability (FO IOP).....	6
Thomas Damm & Thomas Heß	
Improving Climb Performance Prediction in Air Traffic Control with Machine Learning and Full Flight Simulator Verification	10
Dr. Matthias Poppe, Dr. Thomas Pütz & Roland Scharff	
Tactical inputs for a stripless ATC system via Multi-Touch Gestures	21
Dr. Konrad Hagemann & Andreas Udovic	
Erkenntnisse aus dem Projekt “StayCentered - Methodenbasis eines Assistenzsystems für Centerlotsen“	28
Jörg Buxbaum	
Impressum	35

Die Autoren sind, soweit nicht anders gekennzeichnet, Mitarbeiterinnen und Mitarbeiter der DFS. Die Rechte an den Artikeln liegen bei den jeweiligen Autoren.

Übersicht der Artikel in diesem Heft

Thomas Damm & Thomas Heß: Die Zukunft des Flugdatenaustausches: Flight Object Interoperability (FO IOP):

URL: <http://d-nb.info/1197937013/34>, Persistent Identifier URN: URN: urn:nbn:de:101:1-2019102513120867719607.

Hinter dem Begriff „Flight Object Interoperability“ verbirgt sich ein zentrales Element internationaler Bestrebungen wie sie derzeit beispielsweise im europäischen SESAR 2020 Programm stattfinden. Interoperability bezeichnet die Fähigkeit verschiedener Systeme oder deren Komponenten, Informationen untereinander auszutauschen und diese zu nutzen. Der Artikel startet mit einer Begriffsklärung und beschreibt dann nach einem kurzen Abriss der Historie bisheriger Entwicklungsarbeiten die Ergebnisse einer im April 2019 über mehrere Länder verteilt stattgefundenen Simulationsübung. Hierbei waren zahlreiche Flugsicherungsorganisationen (DFS, DSNA, ENAV, Eurocontrol Maastricht, NATS, NAVIAR und PANSO) als auch mehrerer Systemhersteller (Indra, Leonardo und Thales) beteiligt. Die Ergebnisse sind vielversprechend, geben aber auch Hinweise auf weiteren Forschungs- und Entwicklungsbedarf. Basierend auf den Erkenntnissen geben die Autoren einen Ausblick auf weitere Aktivitäten wie beispielsweise eine für Mai 2020 geplante zweite Simulationsübung, sowie vorgesehener Aktivitäten zur Standardisierung.

Dr. Matthias Poppe, Dr. Thomas Pütz & Roland Scharff: Improving Climb Performance Prediction in Air Traffic Control with Machine Learning and Full Flight Simulator Verification.

URL: <http://d-nb.info/1197928634/34>, Persistent Identifier URN: URN: urn:nbn:de:101:1-2019102511135720523136.

Diese Veröffentlichung beschreibt die Ergebnisse einer Studie zum maschinellen Lernen und zeigt einen möglichen Anwendungsfall bei der Flugbahnvorhersage in der Flugsicherung. In der Studie wurde ein neuronales Netz verwendet, um Flughöhen für einen Zeitraum von bis zu 6 Minuten vorherzusagen. Dieses neuronale Netz wurde mit repräsentativen Merkmalen gespeist, die zuvor aus operationellen Mode-S-Daten abgeleitet wurden. In zertifizierten Full Flight Simulatoren wurden parallel dazu Vergleichsdaten erzeugt. Hierbei wurden Steigprofile vom Takeoff bis zum Erreichen des Reisefluges (Top of Climb) unter Variation bestimmter aufgezeichnet. Abschließend wurden die Vorhersagen der Flughöhen auf Basis der operationellen Daten einerseits mit den Vorhersagen auf Basis der Simulatordaten andererseits verglichen. Die kontrollierten Bedingungen des simulierten Fluges erlaubten Rückschlüsse auf den Flugverkehr im operationellen Alltag. Darüber hinaus konnten durch die Simulatordaten weitere repräsentative Merkmale zur Verbesserung der Flugbahnvorhersage mittels neuronaler Netze gewonnen werden.

Dr. Konrad Hagemann & Andreas Udovic: Tactical inputs for a stripless ATC system via Multi-Touch Gestures.

URL: <http://d-nb.info/1190060892/34>, Persistent Identifier URN: URN: urn:nbn:de:101:1-2019070814023078574109.

Deutsche Zusammenfassung: Der Artikel beschreibt die Ergebnisse einer Technischen Demonstration, die im Rahmen des europäischen SESAR Projektes „Workstation, Controller Productivity“ im Zeitraum von 2016 - 2019 bei der DFS zusammen mit dem Hersteller INDRA durchgeführt wurde. Ziel der Technischen Demonstration war die Darstellung der konzeptionellen und technischen Machbarkeit eines auf Multi-Touch Gesteneingaben basierenden User Interface Konzepts für häufige, taktische Eingabeerfordernisse am Lotsenarbeitsplatz. Das Konzept wurde in enger Zusammenarbeit mit der DFS Niederlassung München, Lotsen weiterer DFS Niederlassungen, sowie der Hochschule Osnabrück prototypisch vom DFS Systemhaus und Forschungszentrum entwickelt und von INDRA in einen industriebasierten Prototypen (iCAS) überführt.

Jörg Buxbaum: Erkenntnisse aus dem Projekt „StayCentered - Methodenbasis eines Assistenzsystems für Centerlotsen“.

URL: <http://d-nb.info/1201390273/34>, Persistent Identifier URN: URN: urn:nbn:de:101:1-2019121015383104571833

Im BMBF-Verbundprojekt „StayCentered – Methodenbasis eines Assistenzsystems für Centerlotsen“ hat die Technische Universität Chemnitz Ansätze für eine Emotionserkennung bei Fluglotsen entwickelt und auf ihre Einsatzmöglichkeiten zur Zustandserfassung hin untersucht. In diesem Artikel werden die Erkenntnisse aus den Projektberichten zusammengefasst, eine individuelle Bewertung der Versuchsansätze vorgenommen sowie ein Resümee aus Sicht der Flugsicherung gezogen.

Editorial

Liebe Leserinnen, liebe Leser,



seit vielen Jahren erscheint die DFS Forschungszeitschrift „Innovation im Fokus“ bislang zweimal jährlich sowohl als Druckausgabe, wie auch als elektronische PDF-Variante auf unserer DFS Webseite.

Mit dieser Art der Herausgabe war verbunden, dass die Autorinnen und Autoren für ihre Beiträge fest vorgegebene Fertigstellungstermine einhalten mussten, was aufgrund der vielfältigen, häufig ebenfalls unter Termindruck stehenden Arbeitsaufgaben als zunehmend schwierig und in Zeiten der Digitalisierung auch nicht mehr zeitgemäß erschien. Zudem konnte es aufgrund dieses Veröffentlichungsrhythmus sein, dass Artikel erst viele Wochen nach Fertigstellung tatsächlich erscheinen konnten.

Aus diesen Gründen haben wir uns entschieden die einzelnen Beiträge von „Innovation im Fokus“ zukünftig unmittelbar nach ihrer Fertigstellung und unabhängig von Produktionsterminen einer Ausgabe vorab online auf der DFS-Webseite zu veröffentlichen. Mit der Veröffentlichung geht jetzt die zusätzliche Einzelregistrierung eines jeden Artikels in der Deutschen Nationalbibliothek einher. Dies entspricht einer Praxis, die bei vielen Fachzeitschriften und Journals bereits seit Langem zum Standard gehört.

Auch bezüglich der Gesamtausgaben haben wir uns für eine neue, den Erfordernissen der Zeit angepasste Vorgehensweise mit einer optionalen Halbjahresausgabe und einer fixen Ausgabe zum Jahresende entschieden. Das bedeutet, sofern zur Jahresmitte ausreichend Artikel eingereicht wurden, werden diese wie gewohnt als Halbjahresausgabe zusammengefasst veröffentlicht. Andernfalls werden alle Beiträge eines Jahrgangs am Jahresende in einer Gesamtausgabe „Innovation im Fokus“ sowohl als Printausgabe wie auch als online-Sammelband auf der DFS-Webseite unter <http://forschungszeitschrift.dfs.de/> publiziert.

Wir hoffen Sie bleiben uns auch weiterhin gewogen.

Ihr Günther Achatz
(Herausgeber Innovation im Fokus)

Die Zukunft des Flugdatenaustausches: Flight Object Interoperability (FO IOP)

Thomas Damm & Thomas Heß

Einleitung

Im April 2019 fand im Rahmen des SESAR 2020 Projektes PJ.18 „4D Trajectory Management“ eine Simulationsübung statt. Thema der Übung war die sogenannte „Flight Object Interoperability“ (FO IOP). Unter Interoperabilität versteht man die Fähigkeit von verschiedenen Systemen oder deren Komponenten, Informationen untereinander auszutauschen und diese zu nutzen. Beteiligt waren neben den Flugsicherungsorganisationen DFS, DSNA, ENAV und Eurocontrol Maastricht auch NATS, NAVIAR und PANSO. Zudem wirkten die Flugsicherungs-systemhersteller Indra, Leonardo und Thales mit. Durch diese Übung konnte nachgewiesen werden, dass europäische Kontrollzentralen des oberen Luftraums in der Lage sind, flugrelevante Daten in Form eines sogenannten Flight Object in Echtzeit miteinander auszutauschen.

Was ist FO IOP?

FO IOP soll den Flugdatenaustausch mittels OLDI (On-Line Data Interchange) ersetzen und neue Verfahren unterstützen. Dazu zählt die nahtlose Koordination von Ein- und Ausflugbedingungen und die Kontrollübergabe zwischen Kontrollzentralen. FO IOP ermöglicht den von einem Flug betroffenen Parteien am Boden, flugrelevante Informationen zu lesen, zu aktualisieren und untereinander auszutauschen sowie abzustimmen. Zu den betroffenen Parteien am Boden zählen Air Navigation Service Provider (ANSP), Network Manager (NM), Aircraft Operations Center (AOC) und Flughafenbetreiber. Mit FO IOP wird unter anderem erreicht, dass sich alle an einem Flug beteiligten Lotsen in Europa zu jeder Zeit ein einheitliches, aktuelles Bild des Flugstatus machen können. Ein Merkmal von FO IOP ist, dass Systemfunktionen, die bisher nur innerhalb einer Kontrollzentrale möglich waren, dann auch mit Kontrollzentralen anderer Flugsicherungsorganisationen (bzw. Systemhersteller) möglich werden. Beispiele wären der Austausch eines Release, der Geschwindigkeit und der Route - sowohl upstream als auch downstream.

Die Datenstruktur eines Flight Object enthält mehrere sogenannte Cluster, in denen die Daten themenbezogen zusammengefasst werden. Wichtige Cluster sind beispielsweise das sogenannte Flight Script und die IOP Trajektorie (Trajectory, siehe Abbildung 1).

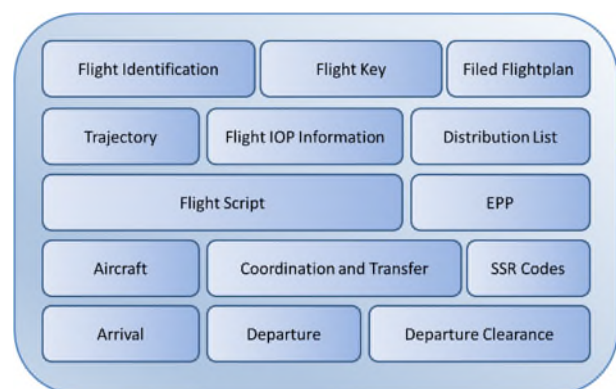


Abbildung 1: Cluster eines Flight Object.

Das Flight Script besteht hauptsächlich aus

- der erweiterten Flugroute (als Folge von 2D Punkten)
- angewiesenen Beschränkungen wie Geschwindigkeit, Ankunftszeiten (CTA) und Hold
- Freigaben

Die IOP Trajektorie (als Folge von 4D Punkten) wird durch den Flight Object Manager derjenigen Kontrollzentrale berechnet, die den Flug aktuell unter Kontrolle hat. Solange eine Kontrollzentrale einen Flug kontrolliert, nimmt diese die Rolle des Flight Data Manager/Publisher (FDMP) ein. Somit ist sie alleinig für das direkte Ändern und Veröffentlichen des Flight Object verantwortlich. Kontrollzentralen, deren Zuständigkeitsbereich ein Flug kreuzen wird, werden als Flight Data Contributor (FDC) bezeichnet. FDC können Änderungen am Flight Object über den FDMP veranlassen. Der Network Manager (NM) fungiert in der Planungsphase eines Fluges als FDMP. Der NM versendet das erste Flight Object, welches zukünftig den Initial Flight Plan (IFPL) ersetzt, an die FO IOP ertüchtigten Kontrollzentralen. Während der Flugphase nimmt der NM die Rolle eines FDC ein.

Andere Instanzen, die berechtigtes Interesse an den Daten eines Flight Object haben, können als Flight Data User (FDU) lesend zugreifen. Dies wären beispielsweise Kontrollzentralen, durch deren Area of Interest ein Flug verläuft. Mit geändertem Flugverlauf können sich die Rollen ändern. Beispielsweise bei Änderung der Sektorsequenz durch eine Routenänderung wird eine FDU zu einem FDC, weil der Flug nun, anders als ursprünglich geplant, den Zuständigkeitsbereich kreuzt.

Als Basistechnologie für FO IOP dient das systemweite Informationsmanagement (SWIM). Grundsätzlich stehen mehrere SWIM-Profilen zur Verfügung. Für FO IOP wird das sog. SWIM Blue Profile angewendet, welches auf Webservices (Request-Response) und Data Distribution Services (Publish-Subscribe) beruht. Das Profil unterstützt Realzeitanwendungen sowie gesicherte Datenübertragungen und ist hochverfügbar.

Historie

Beginnend im Jahre 2002 wurden in Europa mehrere Studien von Eurocontrol, Coflight und iTEC zum Thema FO IOP durchgeführt. Die EUROCAE Arbeitsgruppe WG-59 nutzte deren Ergebnisse für die Erstellung des Standards ED-133 „Flight Object Interoperability Specification“ [1], den sie 2009 veröffentlichte.

Von 2010 bis 2016 wurde FO IOP in verschiedenen SESAR 1 Projekten in Prototypen umgesetzt und validiert. Die Validierungsübungen erreichten den Reifegrad V2 (gemäß E-OCVM [2]) bzw. TRL4 (NASA Technology Readiness Level [3]).

2014 veröffentlichte die Europäische Kommission die DVO 716/2014 „über die Einrichtung des gemeinsamen Pilotvorhabens für die Unterstützung der Durchführung des europäischen Masterplans für das Flugverkehrsmanagement“ (PCP-IR) [4]. Darin wird unter anderem ein erster Informationsaustausch über Flugwege (i4D) zwischen europäischen ATS-Stellen mittels eines ersten systemweiten Informationsmanagements (iSWIM) gefordert.

Um diese Forderungen umzusetzen, arbeiten die europäischen Partner seit 2017 in der SESAR 2020 Solution 18-02b „FO IOP“ zusammen. Im Rahmen der Solution werden die operationellen und technischen Anforderungen erstellt und zwei technische Validierungsübungen geplant und durchgeführt.

SESAR 2020 Validierungsübung zu FO IOP

Die erste FO IOP Validierungsübung in SESAR 2020 fand auf einer, über Langen (Karlsruhe UAC), Maastricht

(Maastricht UAC), Toulouse (Reims UAC, Genf & Zürich UAC) und Rom (Padua & Mailand UAC) verteilten Simulationsumgebung statt (siehe auch Abbildung 2). Es handelte sich um eine Realzeitsimulation für den oberen Luftraum. Die Verkehrsszenarien wurden auf Basis von aufgezeichnetem Verkehr erstellt. Operationelle Experten und Fluglotsen (für die DFS aus Karlsruhe und München) kontrollierten, unterstützt von Simulationspiloten, den Verkehr an Systemprototypen (DFS: iCAS Phase 2 Prototyp). Dazu dienten in der Solution beschriebene Anwendungsfälle zu Themen wie „Coordination and transfer“, „Management of the FO flight script“ und „Informative distribution between system instances“. In der folgenden Abbildung ist der simulierte Luftraum dargestellt:

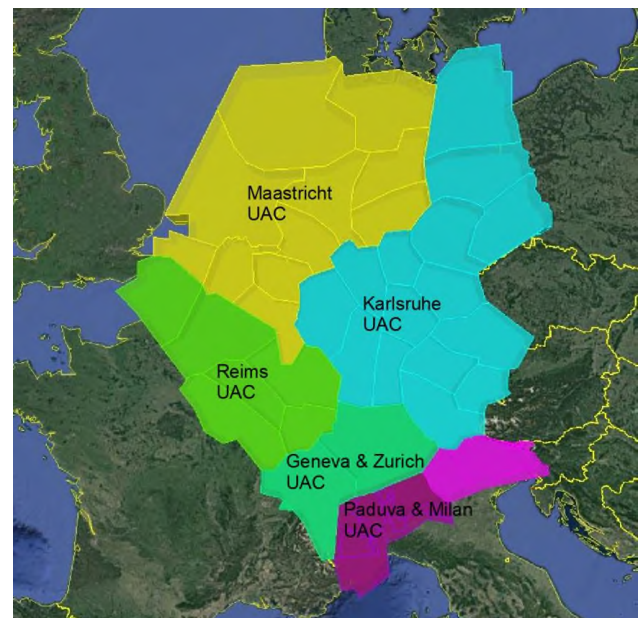


Abbildung 2: Grafische Darstellung des simulierten IOP Luftraums.

Die räumlich verteilten Systemprototypen wurden jeweils lokal durch einen Simulator (DFS: NEWSIM/ ATGateway/ Shared Virtual Sky (SVS) Gateway) mit Radar- und Flugplandaten stimuliert. Der Simulationsstart und die von den Simulatoren erzeugten Positionsdaten der Flugzeuge wurden mit SVS über das Internet synchronisiert. Die Systemprototypen tauschten das Flight Object via SWIM Knoten über den Pan-European Network Service (PENS) aus (siehe Abbildung 3).

Bei SVS handelt es sich, ebenso wie bei ATGateway, um eine Implementierung des EUROCAE-Standards ED-147 [5]. Dieser wird von der EUROCAE Arbeitsgruppe WG-81 gepflegt und beschäftigt sich mit der Interoperabilität von ATM-Validierungsplattformen.



Durch die erste Validierung konnten unter anderem folgende Problemfelder erkannt werden:

Nichtsdestotrotz sehen die operationellen Teilnehmer folgendes Potential durch die Einführung von FO IOP:

Verbessertes Situationsbewusstsein durch eine ständig aktualisierte, konsistente Flugplanung

Verbesserte Werkzeuge zur Konflikterkennung und -lösung

Ausblick

Auf Basis der Ergebnisse aus der SESAR 2020 Solution 18-02b wird der Standard ED-133 in der EUROCAE Arbeitsgruppe WG-59 überarbeitet werden. Zunächst werden die Beteiligten voraussichtlich Ende 2019 bis Anfang 2020 einen inoffiziellen Entwurf erstellen. Mitte 2021 soll dann ED-133 Revision A veröffentlicht werden. Der ED-133 Entwurf dient als Grundlage für die IOP Industrialisierung. Die DFS plant dies gemeinsam mit unseren iTEC Partnern und dem Maastricht UAC ab 2020.

Die Umsetzung von FO IOP in den Betrieb wird maßgeblich durch den SESAR Deployment Manager koordiniert. Gemäß der aktuell geltenden DVO 716/2014 (PCP-IR) sollen iSWIM und i4D, und somit FO IOP, zum 1. Januar 2025 angewendet bzw. ermöglicht werden. Der genaue Zeitplan der Implementierung und Möglichkeiten der Förderung werden momentan noch von der Europäischen Kommission evaluiert. In der IOP Roadmap wird davon ausgegangen, dass FO IOP in Europa bis Sommer 2028 voll betriebsbereit (FOC) ist.

Dieser Artikel berichtet zu Daten, die im Rahmen eines vom SESAR Joint Undertaking geförderten Projektes (Grant Agreement Nummer 724161) im Horizon 2020 Forschungs- und Entwicklungsprogramm der Europäischen Union erhoben wurden. Dieser Artikel stellt nur die Sichtweise der Autoren dar. Das SESAR JU ist nicht verantwortlich für jegliche Art der Weiterverwendung von Informationen, die dieser Artikel enthält.

ANSP	Air Navigation Service Provider; Flugsicherungsorganisation
AOC	Aircraft Operations Center
CTA	Calculated Time of Arrival
DSNA	Direction des Services de la Navigation Aérienne (Französische Flugsicherung)
DVO	Durchführungsverordnung
ENAV	Società Nazionale per l'Assistenza al Volo (Italienische Flugsicherung)
E-OCVM	European Operational Concept Validation Methodology [2]
EPP	Extended Projected Profile

EUROCAE	European Organisation for Civil Aviation Equipment
FDC	Flight Data Contributor
FDMP	Flight Data Manager/Publisher
FDU	Flight Data Unit
FO	Flight Object
FO IOP	Flight Object Interoperability
FOC	Full Operational Capability
i4D	Initial 4D (Trajectory Management)
IFPL	Initial Flight Plan
iSWIM	Initial System Wide Information Management
ITEC	Interoperability Through European Collaboration
NASA	National Aeronautics and Space Administration
NATS	National Air Traffic Services (Britische Flugsicherung)
NAVIAR	Navigation Via Air (Dänische Flugsicherung)
OLDI	On-Line Data Interchange
PANSA	Polish Air Navigation Services Agency (Polnische Flugsicherung)
PCP-IR	Pilot Common Project Implementing Rule [4]
PENS	Pan-European Network Service
SVS	Shared Virtual Sky (Interoperabilitätsservice für Simulatoren von Sopra Steria)
SWIM	Systemweites Informationsmanagement
TRL	Technology Readiness Level [3]
UAC	Upper Area Control Center

Referenzen

- [1] EUROCAE (2009). Flight Object Interoperability Specification. ED-133.
- [2] Eurocontrol (2010). European Operational Concept Validation Methodology (EOCVM). URL aufgerufen am 16.9.2019: <https://www.eurocontrol.int/publication/european-operational-concept-validation-methodology-eocvm>
- [3] NASA (28.10.2012). Technology Readiness Level. URL aufgerufen am 16.9.2019: https://www.nasa.gov/directorates/heo/scan/engineering/technology/txt_accordion1.html
- [4] Europäische Kommission (2014). Einrichtung des gemeinsamen Pilotvorhabens für die Unterstützung der Durchführung des europäischen Masterplans für das Flugverkehrsmanagement. Durchführungsverordnung Nr. 716/2014.
- [5] EUROCAE (2017). ATM Validation Platforms Interoperability Specification. ED-147A.

Improving Climb Performance Prediction in Air Traffic Control with Machine Learning and Full Flight Simulator Verification

Dr. Matthias Poppe, Dr. Thomas Pütz & Roland Scharff

Deutsche Zusammenfassung

Diese Veröffentlichung beschreibt die Ergebnisse einer Studie zum maschinellen Lernen und zeigt einen möglichen Anwendungsfall bei der Flugbahnvorhersage (Trajektorien-Prädiktion) in der Flugsicherung. In der Studie wurde ein neuronales Netz (hier: Deep-Feedforward-Netzwerk) verwendet, um Flughöhen für einen Zeitraum von bis zu 6 Minuten vorherzusagen. Dieses neuronale Netz wurde mit repräsentativen Merkmalen (Features) gespeist, die zuvor aus operationellen Mode-S-Daten (Enhanced Surveillance) abgeleitet wurden. In zertifizierten Full Flight Simulatoren (Airbus A340 und A319) wurden parallel dazu Vergleichsdaten erzeugt. Hierbei wurden Steigprofile vom Takeoff bis zum Erreichen des Reisefluges (Top of Climb) unter Variation bestimmter Parameter, wie z.B. Takeoff-Gewicht oder Gegen- und Rückenwind-Bedingungen, aufgezeichnet. Analog zu den operationellen Daten wurden die im Simulator gewonnenen Merkmale in das neuronale Netzwerk eingespeist, um hier ebenfalls eine Vorhersage zu berechnen. Abschließend wurden die Vorhersagen der Flughöhen auf Basis der operationellen Daten einerseits mit den Vorhersagen auf Basis der Simulatordaten andererseits verglichen. Die kontrollierten Bedingungen des simulierten Fluges erlaubten Rückschlüsse auf den Flugverkehr im operationellen Alltag. Darüber hinaus konnten durch die Simulatordaten weitere repräsentative Merkmale zur Verbesserung der Flugbahnvorhersage mittels neuronaler Netze gewonnen werden.

Abstract

A deep feedforward network has been used to predict the flight level with look ahead time up to six minutes for climbing flights. Representing features were developed from operational Mode S Enhanced Surveillance data. In parallel, using certified A340 and A319 Full Flight Simulators, complete climb profiles from Take Off to Top of Climb have been recorded. Certain parameters were modified, e.g. Take Off weight or head- and tailwind conditions. The features from the Full Flight Simulator were fed to the neural network which has been trained with the operational data. The predictions of the flight level in both cases are compared. Because the

parameters in the simulator were known and controlled, it allows us to infer the settings of the operational traffic and to improve the features by adding wind information and recognizing the acceleration phase.

Keywords: Air Traffic Control, trajectory prediction, machine learning, neural network, full flight simulator

Introduction

The accurate prediction of the aircraft trajectory is a cornerstone for the majority of current and future operational concepts [1] and controller decisions support tools. Of high interest is the improvement of the vertical part of the trajectory because of the bandwidth of vertical climb rates, speed development and the associated operational and technical uncertainty [2]. Some factors (not exhaustive) with high impact to the vertical rate are the aircraft weight and cost index, the atmospheric conditions, operational restrictions and human intervention like controller clearances and pilot behavior.

In the last years, many research activities regarding trajectory prediction (TP) rely on parametric models. Typical aircraft characteristics are derived from the EUROCONTROL BADA model [3]. [4] provides an overview of TP research studies. The majority uses a point-mass model with the problem of estimating the a priori not known input data, e.g. Take Off weight. Another unknown is the operational environment of the flight, which may impose other constraints that impact the performance of the aircraft [5].

Current research tries to overcome some of these problems by applying machine learning techniques to estimate some of the most important parameters, e.g. the Calibrated Airspeed [6] or the Take Off aircraft mass [7]. In [8], a hybrid system based on BADA has been developed in an online context where known altitudes were used to fit the model parameters and to predict the remaining part of the trajectory for the climb phase.

Non-parametric approaches are less frequent because they require a huge amount of input data for sufficient training. In [9], a single hidden layer feed forward network was trained for a selected aircraft type. One disadvantage of the non-parametric trajectory prediction

is the difficulty to generalize the model once it has been trained.

This paper describes the results obtained from a deep feedforward neural network, that was used to predict the aircraft's altitude immediately after Take Off up to Flight Level 285 (1FL = 100ft). Supervised learning was applied with four months of recorded operational data from all over Germany. The results were published in [10]. For all aircraft types and looking at major German airports like Frankfurt (EDDF), Munich (EDDM) or Düsseldorf (EDDL) with a prediction horizon of six minutes, the mean absolute prediction error was 8.14 Flight Level (FL) with a standard deviation of 7.21 FL. The advantage of this approach is the rather good generalization because the results are in general not constrained to a specific aircraft type or operational ATC/airport environment. The Feedforward network is able to 'learn' and differentiate the departures from different airports.

In order to further improve the prediction accuracy and to learn more about the chosen input data (features), several climbing flights were conducted in a Full Flight Simulator. In this controlled and almost realistic environment, a variety of parameters have been modified: cost index, Take Off weight and wind. The derived features were also fed into the deep feedforward network that has been trained previously with the data from the operational traffic. A comparison of the predicted vertical profile allows us to infer the most common parameter settings.

After a general introduction to neural networks in the following section, the derived feature vectors for the Machine Learning approach are explained. Then the Full Flight Simulator trials and their analysis are described. Finally, the results of the comparison are presented.

Supervised Learning

Brief Introduction

A thorough introduction to the subject of machine learning with deep feedforward networks can be found e.g. in [11].

The main objective for supervised learning is to train a model with labeled training data. This trained model allows to make predictions of future yet unknown input data (test data). The training data are often called 'features' X of an observed phenomenon (the explanatory or independent variables) while the output data are called 'labels' y . We assume that there exists a relation $y = f(X)$ between X and a target variable y . In

general, a model h is learned with examples of the outputs $(y_0, \dots, y_N)$ associated with the inputs $X^T = (x_0, \dots, x_N)$, approximating f . A performance measure P will be applied to measure the abilities of the machine learning algorithm. For our prediction task, we applied supervised learning with an output scalar $\hat{y}_i \equiv \text{predicted FL}$.

Supervised learning itself can be broadly divided into two different main categories: classification and regression. In classification tasks, the output y belongs to a category or group. For example, we could forecast the aircraft type given certain input trajectories to train a network. In regression tasks, y is a variable with a 'real' numerical value or a vector of values. We will use regression techniques to forecast the flight level, expressed as a numerical value, e.g. FL264.

Feedforward Network

Figure 1 shows a feedforward network with the training cycle of one input layer $\mathbf{x}$, hidden layers $\mathbf{h}^n, \mathbf{h}^k$ and the network output layer $\hat{\mathbf{y}}_i$.

The feedforward network defines the mapping of the supervised target $\mathbf{y} = f(\mathbf{x}; \boldsymbol{\theta})$ to $\hat{\mathbf{y}}_i$ and learns the value of the parameters $\boldsymbol{\theta}$ that result in the best function approximation. $\boldsymbol{\theta}$ may consist of weights w and a bias b . As example, layer k computes an output vector $\mathbf{h}^k$ using the output $\mathbf{h}^n$ of the previous layer, starting with the input $\mathbf{x} = \mathbf{h}^0$. Thus:

$$\mathbf{h}^k = \sigma(\mathbf{b}^n + W^k \mathbf{h}^n) \quad (1)$$

with W^k as a matrix of weights in layer k and σ as the in general non-linear activation function.

The rectified activation function is the default function recommended for use with most feedforward neural networks [11]:

$$\sigma(u) = \begin{cases} 0, & u < 0 \\ u, & u \geq 0 \end{cases} \quad (2)$$

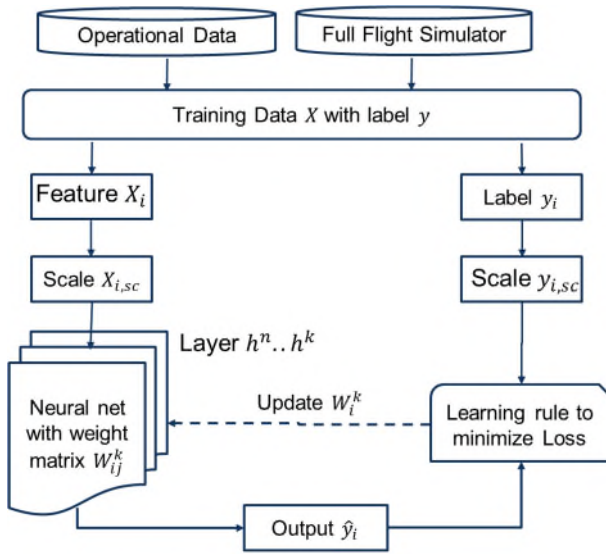


Figure 1: Model Training for Supervised Learning.

The goal is to minimize the cost function or the loss of the model with a specific learning rule. In most cases, a parametric model defines a distribution $p(\mathbf{y}|\mathbf{x};\boldsymbol{\theta})$ and we can use the principle of maximum likelihood, i.e. we use $-\log p(\mathbf{y}|\mathbf{x};\boldsymbol{\theta})$ as cost function. Other cost functions can be used as well, e.g. predicting the median value of $\mathbf{y}$ for each $\mathbf{x}$ (mean absolute error). In optimizing the cost function, it is expected that the chosen performance measure P is optimized as well.

As the parametric model distribution p depends on the features X_i which are directly derived from the raw recorded input data, this paper will show the impact of the derived features and how they can be tuned in order to improve the prediction, using the same feedforward network.

Raw Data and Features

The input data were restricted to easily available transponder Mode S Enhanced Surveillance data [12] that could be obtained from the majority of aircraft (more than 98% in Germany for IFR traffic, according to DFS internal statistics from 2017). The second source of data are the recorded flight data from the Full Flight Simulator with a sample rate of one second.

The raw input data are pre-processed in order to obtain a representative feature vector $\mathbf{X} \in \mathbf{R}^{12}$. In order to determine which features are predictive for the dynamics of the climbing phase, information from multiple sources were gathered, including the knowledge of controllers, pilots and statistics of the given datasets [14][15]. All features are calculated using the Mode S EHS data and the aircraft type [16].

The motivation for the chosen features was to keep the Machine Learning network relatively small compared to an unfiltered set of features with less predictive characteristics. The training is also expected to be faster and more efficient [11].

Decoding of Mode-S EHS

The selected Comm-B Data Selector (BDS) transponder register are (for further details, please refer to [13]):

1. BDS 4.0 - Selected Vertical Intention
2. BDS 5.0 - Track and Turn Report
3. BDS 6.0 - Heading and Speed Report

The flight example in Figure 2 shows some extracted parameters and how they evolve over time. The IAS is shown in red color with the restriction of 250 knots below FL100. The acceleration phases exhibit the typical drop in the vertical rate (orange line). The Flight Control Unit (FCU) Selected Altitude (blue line) reflects the current clearance. The figure also shows a typical initial peak in the climb rate shortly after Take Off (≈ 3000 feet/minute).

The two vertical bars mark the prediction interval: the earliest predictions are made 60 seconds after take-off while the latest predictions are made not later than Flight Level 220. The main reason for this choice is to avoid making forecasts, once the aircraft has been leveled off.

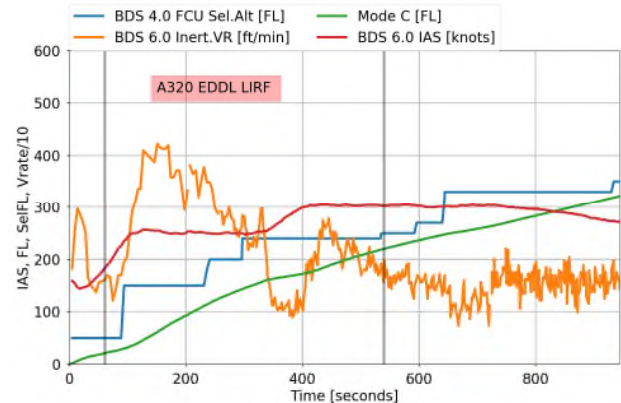


Figure 2: Example of obtained Mode S flight data.

Being interested in making predictions for climbing flights, a flight was defined to be “climbing” whenever the recorded data fulfills:

$$\min_t \{FL(t)\} < 20, \max_t \{FL(t)\} > 285 \quad (3)$$

In total, more than 400.000 climbing flights all over Germany were extracted in spring and summer 2018. From the Full Flight Simulator, we recorded data from 27 runs.

Labels Extraction and Scaling

The aim is to predict the flight level of the aircraft up to the next six minutes (360s). The label for each feature vector contains the FL values in $n * 90$ seconds with $n = \{1, 2, 3, 4\}$. The network was trained separately for each FL prediction horizon.

For each flight, random indices were chosen and the feature vector as well as the corresponding labels were created. In total, close to 400.000 features and labels were generated leading to a feature vector of shape $X(397.242, 12)$. The data were randomly split with shuffling into 80% training data and 20% test data.

In a second step, the features and labels were standardized by removing the mean and scaled them to unit variance. Centering and scaling happen independently on each feature by computing the relevant statistics on the samples. Standardization is a common requirement for many estimators. If a feature has a variance that is orders of magnitude larger than others, it might dominate the objective function and make the estimator unable to learn from other features correctly as expected [17].

Features

The features are shown in Table I. They were derived from the Mode S EHS data, as well from the operational data as from the flight simulator. They have been modified in comparison to [10] in order to better reflect the acceleration phase (x_2 and x_4).

TABLE I. FEATURES

	Feature	Unit
x_0	Take Off Safety Speed V_2	[knots]
x_1	Actual Rate of Climb ROC	[feet/minute]
x_2	$\Delta ROC / \Delta t$	[feet/minute/second]
x_3	Initial ROC Peak	[feet/minute]
x_4	$\Delta IAS / \Delta t$	[knots/second]
x_5	Heading Change Flag	[bool]
x_6	Altitude Restriction Flag	[bool]
x_7	Initial heading	[degree]
x_8	Current FL	[flight level]
x_9	Current IAS	[knots]
x_{10}	Aircraft Type	-
x_{11}	Departure Airport	-

Full Flight Simulator Verification

Set-up

For the verification of the correct algorithm behavior and in order to exclude any unknown influence factors, a full flight simulator evaluation was performed.

Several full flight simulator sessions were executed using different simulator types and different settings.

All simulators were Level D flight synthetic training devices qualified by the national aviation authority in Germany, LBA. Provider of the training infrastructure and the know-how was Lufthansa Aviation Training (LAT, part of Lufthansa AG) in their facilities in Frankfurt. Chair of flight guidance and air transportation of Technische Universität Berlin (TUB) was supporting the activity with their know-how.

The following simulators were used:

- Airbus A340-300 (Thales, CFM56-5C4)
- Airbus A319-100 (Thales, CFM56-5A5)

These types were selected to cover long range as well as short and medium range operations (refer to Figure 3 for cockpit view of A340-300 simulator).



Figure 3: Cockpit View A340-300 Simulator [source: Lufthansa Aviation Training].

Several series of simulation runs had been prepared. Overall there were four simulation sessions with approximately 6-8 runs per session.

Basic requirements for all runs and sessions had been determined. Several typical flights for airline operations have been selected. The focus was on the departure procedure. The flights were executed according to flight crew operations manual without ATC intervention and also without manual control as well as selected auto flight guidance modes (except it was required by the scenario). Manual control was only applied during the

take of procedure as long as required. Furthermore, a most recent navigation database was used. All scenarios are based on a detailed flight planning that includes a fuel and performance calculation with according professional tools.

Each simulation run was recorded and documented. Main source of information was the recording of simulation parameters. In addition, a video recording (refer to Figure 4) and observation notes were taken.

Chair of flight guidance of TUB provided the tool to record flight parameters of interest in real time at a configurable frequency; in our case 1 record per second. Parameters like time, position (latitude/longitude), altitude (Barometric Altitude, Radio Height), speed (Calibrated Airspeed (CAS), True Airspeed (TAS), Mach), Gross Weight (GW), attitude angles (pitch roll, heading), power setting and system switching (gear, flaps, Autopilot A/P etc.) were recorded.



Figure 4: Primary Flight Display and Navigation Display [source: Technical University Berlin].

For each aircraft type a typical routing was assigned. No limitations in the take-off flight path were given. Bleed Air and Anti-Ice settings were kept equal.

The following experimental conditions were varied in order to assess their effect on the observed trajectory. Takeoff weight and performance with speeds as well as power setting (flex power) were varied. The Cost Index (CI) was investigated in different areas of the range from 0 to 99. For both aircraft types, CI=0 indicates a minimum fuel consumption while CI=99 results in a minimum time flight. Wind and weather conditions were selected with focus on strong linear wind fields with a vertical change for increasing altitudes only. Direction and strength (e.g. 070/030 70° with 30 knots) was kept

constant over time for all the layers of the wind model. Load and trim was also kept constant.

Table 1 -Airbus A340 A319 FFS experimental conditions

RUN	OPS PROCEDURE	WIND	GROSS WEIGHT	POWER SETTING FLEX T	COST INDEX
SESSION 1 A343					
#01	TO & CLB	070/30	271 t	F36	99
#02	TO & CLB	070/30	190 t	F54	0
#03	TO & CLB	070/30	190 t	F54	99
#04	TO & CLB	0/0	271 t	F33	99
#05	TO & CLB	070/30	190 t	TOGA	0
SESSION 2 A319					
#06	TO & CLB	0/0	55 t	F58	0
#07	TO & CLB	070/40	68 t	F58	0
#08	TO & CLB	0/0	55 t	F56	65
#09	TO & CLB	250/15	68 t	F48	99
#10	TO & CLB	250/15	68 t	F48	35

Apart from wind the weather setting was a standard atmosphere without any weather phenomena except wind (ISA, CAVOK with NOSIG¹).

The aircraft was guided in managed mode except for the take-off. Configuration changes were kept comparable. They were triggered when reaching the according speeds altitudes for the flap setting or the thrust respectively.

The climb procedure was a Constant IAS/Mach climb with econ speeds with a 250 knots speed limit until FL100.

Main focus was the phase from start of acceleration immediately after Take Off until reaching cruising flight

¹ International Standard Atmosphere, Clouds and Visibility OK, No Significant Change

level. The impact of the take-off procedure on this phase was investigated as well. Due to the takeoff settings and a potentially resulting different takeoff flight path, FL100 with 250kts speed restriction was selected as the starting point of the diagrams in order to better compare the effect on the climb procedure.

It was concluded after some runs that under these controlled experimental conditions, there is only a very small statistical variation. There is no justification for increasing the number of run samples for one set of given parameters in order to allow a conclusion on the observed effect.

Simulator Results

In theory, the main factors for the climb performance are weight, climb speeds, wind, temperature, pressure altitude, drag and available power. Cost index may have an influence on the speed.

In the according data logs the effect of power setting, gross weight, cost index and wind on climb performance was assessed.

A first observation was that the available climb power was constant for the given scenarios and parameter variations after the takeoff procedure. In case all other parameters are equal, especially temperature, the recording do not show any difference for the climb procedure. Therefore, no further comparison can be done here.

A session with an Airbus A340-300 was starting the campaign. As expected, the gross weight has a clear impact on the climb performance. Figure 5 shows the effect of a different gross weight for a low and a high gross weight scenario. The aircraft is reaching its cruising altitude at an earlier point in time. The necessary climb rate is higher.

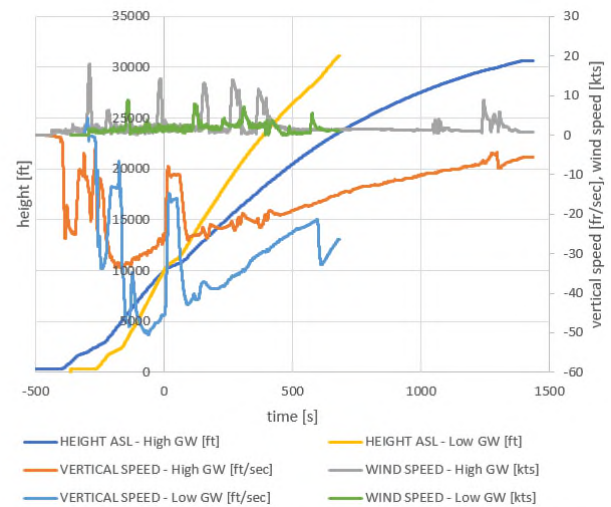


Figure 5: A340: Gross weight influence on vertical rate and profile under similar wind conditions.

A bit surprising and not necessarily representative is that the cost index seems to have no clear effect on climb performance here, at least until a certain altitude that can be seen in Figure 6. Potential reason might be that this is an aircraft specific behavior due to a low climb performance.

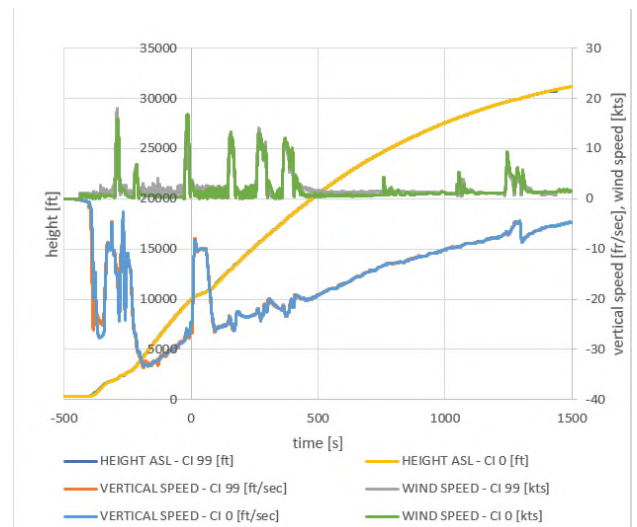


Figure 6: A340: Cost index influence on vertical rate and profile under similar wind conditions.

Looking at the wind in Figure 7, no obvious influence of a steady wind without vertical component could be observed on the first sight. However, a deeper analysis was revealing an effect here.

A steady wind or slowly increasing or decreasing wind has no effect respectively a negligible one, as expected. However, in the first phase of the flight, where wind speed fluctuations could be detected, a slight impact on

the vertical rate is visible. Overall this leads potentially to a very marginal change in the vertical profile. Figure 7 shows the vertical climb profile without wind and with a headwind of 30 knots at ground level up to 70 knots above Flight Level 200.

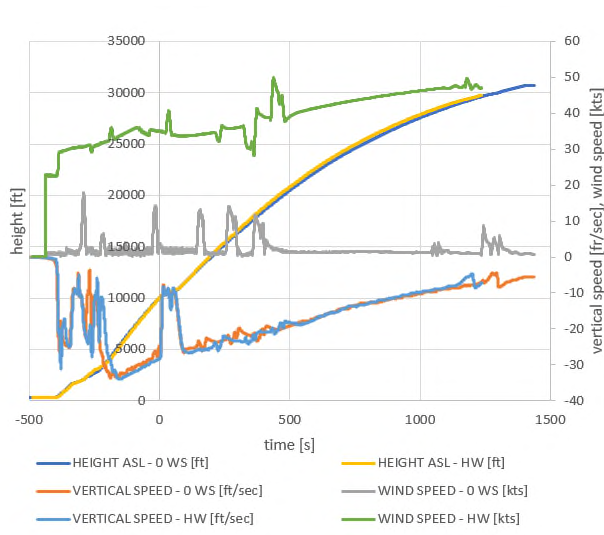


Figure 7: Airbus A340 FFS results with headwind and no wind.

The simulation sessions with the A319-100 show similar behavior mainly for gross weight. Therefore, the weight impact is not shown here.

Concerning the cost index (See Figure 8), for A319 the cost index has an observable impact on the vertical profile. This is fully in line with available documentation and literature (compare e.g. with [15][24])

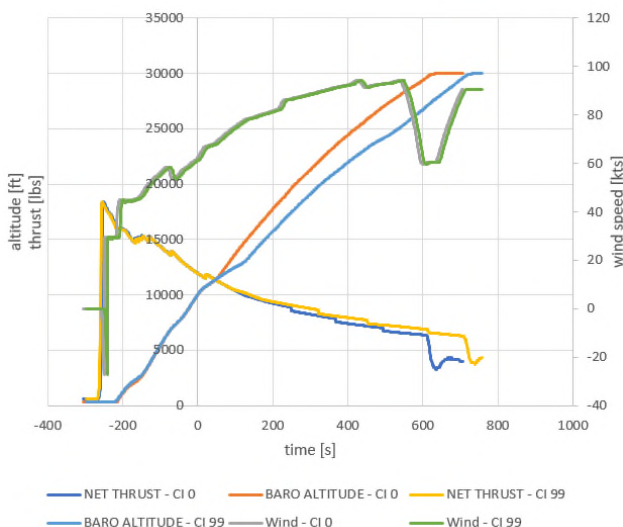


Figure 8: A319: Cost index influence on vertical rate and profile under similar wind conditions.

The wind influence has a stronger effect, compared to the runs with the A340. In the wind scenarios, a rather

strong headwind was introduced from 40 knots at ground level up to 100 knots above FL200. In runs #09 and #10, the maximum allowed tailwind for Take Off at ground level of 15 knots was simulated up to 100 knots tailwind above FL200. Figure 9 shows the resulting climb performance in comparison.

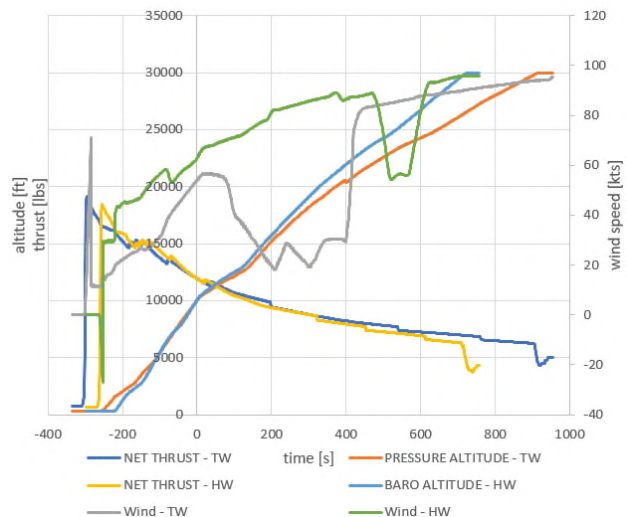


Figure 9: A319 FFS results with head and tailwind scenario.

The blue and orange line show the pressure altitude in the head- and tailwind run respectively. Flight Level 300 was reached with a difference of about 200 seconds with identical flight control settings. Most likely the steep wind gradients increased the effect. It was confirmed that the wind field was static and not changing strength over time and in the same altitude.

First results revealed that vertically increasing or decreasing constant winds influences the climb rate (feet per minute) in a significant way. A head wind that is increasing during climb rises the observed climb rate. A tailwind has the opposite effect. Depending on the behavior of the wind field a degradation, which is the critical influence for ATC, needs to be considered by controllers as well as by support tools. A better knowledge of the atmospheric environment helps to strengthen the accuracy and the performance of the ATM system und the offered service by the ANSP.

Sometimes ATCOs complain about lack of performance which cannot be explained by simple atmospheric conditions and load of cost index only. Approaching strong winds like the stream could have an additional effect here as well.

Effects of wind fields are known in theory and also for gliders. However, in some cases this could be observable also in case of airline operations. Although there are no practical rules for wind gradients in daily

ATC control, controllers have to face this effect in their work. Also flight crew material considers usually a constant wind speed influence only.

The observed effect is well documented in the literature but mainly for light gliders or for approach, see [23].

Following formula explains the effect (formula 6.3.7 from Brockhaus [23]) and shows the relationship between thrust F , weight G , glide ratio γ_{GL} , glide path γ , vertical wind component w_W , flown airspeed V_A , time derivate of the horizontal wind component u_W and gravitational acceleration g .

$$\frac{F}{G} = -\gamma_{GL} + \gamma + \frac{w_W}{V_A} + \frac{u_W}{g} \quad (4)$$

with $w_W = 0$ and F, G, g and $\gamma_{GL} = \text{const.}$

With increasing altitude the time derivative of the horizontal wind component is not zero which leads to a steeper or shallower flight path and an increasing respectively decreasing rate of climb in case the airspeed is kept constant.

Taking these detailed conclusions examples into account, FFS results in general helped to understand which settings are applied in reality. Furthermore this led to conclusions on which features were relevant to further improve the Deep Learning approach and to make it more robust.

Climb Performance Prediction

The neural network has been developed for the prediction of the climb performance with the deep learning library *Keras*, together with *tensorflow* [18][19]. These open source software libraries support flexible machine learning architectures and allow computation across a variety of platforms including Graphic Processor Units (GPU).

Network Training

The obtained results were achieved with a network of six hidden layers while each layer contains a different number of nodes in descending order with one node for the output layer. The rectified linear unit has been chosen as activation function for each layer. Different optimizers were tested (ADAM, Adadelta, stochastic gradient descent (SGD), refer to [21] for an overview) but no significant differences in convergence could be observed. Figure 10 shows an example of the learning curve for training and test data with ADAM. At epochs 15 and 75, there is a step drop in the training data loss and less obvious in the test data. This is due to the reduction of the learning rate because the improvement in loss

falls below a configurable threshold (*Keras ReduceLROnPlateau*). After about 100 epochs no further reduction of the loss could be observed.

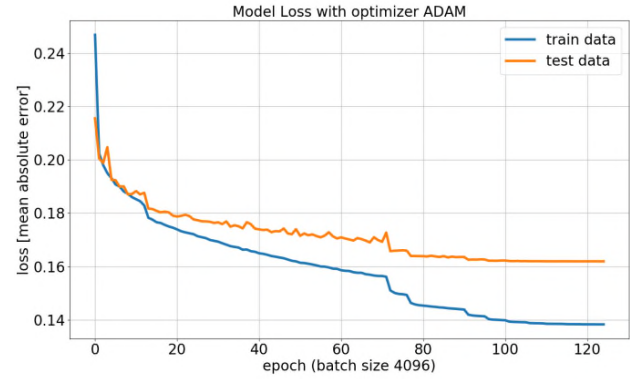


Figure 10: Learning curve for training and test data with new features.

The R^2 statistics provides a measure of fit, i.e. the proportion of variability in y that can be explained using X . An R^2 statistic that is close to 1 indicates that a large proportion of the variability in the response y has been explained by the model [22]. It is defined as:

$$R^2 = 1 - \frac{RSS}{TSS} \quad (5)$$

$$RSS = \sum_i (y_i - \hat{y}_i)^2 \quad (6)$$

$$TSS = \sum_i (y_i - \bar{y}_i)^2 \quad (7)$$

$$\bar{y}_i = \frac{1}{n} \sum_{i=1}^n y_i \quad (8)$$

RSS is the residual sum of squares, TSS the total sum of squares and $\bar{y}_i$ the mean value.

The R^2 value for the A343, obtained from the feedforward network with training and validation data from 12.206 recorded operational flights, is 0.965. This means that almost all variations in the climb behavior can very well explained with the chosen machine learning approach. For the A319, the R^2 value of 0.919 is lower but still close to 1. 204.016 samples were used for training and independent validation (90:10 data split). The lower R^2 value for the A319 aircraft types stems most likely from the fact that this two-engine aircraft type needs higher power settings to compensate for engine failure during Take Off which results in a higher variability of the climb phase.

Results with Data from Full Flight Simulator

TABLE II. and TABLE III. provide for a prediction time of 360 seconds the obtained R^2 score, the mean absolute error *MAE* and the standard deviation *STD DEV* of the error. The *MAE* takes into account the absolute

difference between predicted flight level and the measured flight level (either from operational live data or from the full flight simulator). The MAE is defined as:

$$\frac{1}{n} \cdot \sum_{i=1}^n |y_i - \hat{y}_i| = \frac{1}{n} \cdot \sum_{i=1}^n |e_i| \quad (9)$$

For the Full Flight Simulator runs, the MAE is based on a relatively small number of 16 samples per run. Assuming a standard vertical separation of 1000 feet or 10 FL, we observe a Mean Absolute Error (MAE) of about 5.6 FL for a prediction horizon of six minutes for the A343. The MAE for the A319 is slightly higher with 7.55 FL, possibly caused by the reasons mentioned before. The mean error is less than the vertical separation standard. The standard deviation reflects the uncertainty, to some extent caused by human (controller and pilot) intervention. If the *Managed Mode* flag in the Mode S EHS could be made available, the standard deviation could possibly be decreased. Up to now, only few aircraft (e.g. B787, A350) transmit this flag.

TABLE II. PREDICTION METRICS A343

Data Source	Prediction Time 360 seconds A343				
	MAE	STD DEV	GW	CI	Wind
Live Data with $R^2 = 0.965$	5.60 FL	5.38 FL	?	?	?
FFS Run #01 8	13.91 FL	8.88 FL	271	99	HW
FFS Run #02 9	4.72 FL	2.93 FL	190	0	HW
FFS Run #03 A	6.80 FL	7.49 FL	190	99	HW
FFS Run #04 2	18.91 FL	7.77 FL	271	99	-
FFS Run #05 B	8.27 FL	3.89 FL	190	0	-

HW: head wind; TW: tail wind

TABLE III. PREDICTION METRICS A319

Data Source	Prediction Time 360 seconds A319				
	MAE	STD DEV	GW	CI	Wind
Live Data with $R^2 = 0.919$	7.55 FL	6.41 FL	?	?	?
FFS Run #06 2	3.55 FL	2.51 FL	55	0	-
FFS Run #07 3	6.64 FL	5.62 FL	68	0	-
FFS Run #08 1	5.49 FL	3.87 FL	55	65	HW
FFS Run #09 7	32.44 FL	5.69 FL	68	99	TW
FFS Run #10 8	25.60 FL	6.54 FL	68	35	TW

HW: head wind; TW: tail wind

For the A343 (refer to Table II), best results have been achieved in runs #02, #03 and #05. They have in common a minimum Take Off weight while the cost index has been varied between 0 and 99. In run #05, no flex

power has been used for the Take Off procedure. Head wind was available in runs #01 to #03 between 30 (Ground) and 60 from FL200 onwards) knots. While wind and cost index seem to play a minor role in the prediction capability, the Take Off weight is the most sensitive parameter for our feedforward network. Run #01 and #04 with high CI and maximum Take Off weight show the highest prediction error for this aircraft type.

The picture looks similar for the A319 (refer to Table III). Run #06 without wind and run #07 with headwind (both with CI=0) deliver the best results with a low mean absolute error even below the MAE for the operational live traffic data. The Take Off weight plays for this type a minor role, compared to the A343. Run #08 without wind and medium values for Gross Weight and CI (compared to run #07) shows even less medium prediction errors. In run #09, the level off segment in FL120 of about 60 seconds together with tailwind caused the worst prediction. Also run #10 shows a poor result, most likely caused by the strong tailwind conditions. In reality, it seems unlikely to meet these conditions because 15 knots tailwind is the maximum allowed value for the A319 according to the Electronic Flight Bag (EFB) information.

Conclusion and future work

Deep learning with neural networks has been applied to predict the aircraft climb behavior with a time horizon of six minutes. The achieved accuracy for two selected aircraft types shows a mean absolute error of about 6 to 7 FL, well below the actual vertical separation minimum of 10 FL. The developed features describe the general climb behavior and allow to train the neural network, obtaining for both considered aircraft types statistical R^2 scores above 0.9. The authors conclude that this approach allows to reduce the uncertainty buffers for Decision Support Tools in the climb phase, which is the most challenging part of a flight in terms of trajectory predictions.

The Full Flight Simulator verification has shown that certain aircraft parameters reflect the actually used settings in reality, because the feedforward network was trained with real operational data. For the A343 the gross weight impacts the climb performance which can be deducted from the worse prediction results. The Cost Index seems to have only a minor impact. For the A319, strong tailwind conditions worsen the prediction results, together with level-off clearances which were artificially induced during the simulation runs. The gross weight and also the CI settings seem to play a minor role for this twin-engine aircraft type. While moderate headwind

seems to not impact the climb performance, situation is different for strong wind gradients and strong head- and tailwinds. The feature vector for the neural network should be enhanced accordingly.

The authors also believe that ADS-C with Enhanced Projected Profile (EPP) reports could provide the currently unknown parameters like CI or gross weight which could be included in the features. Additional benefit would be to include the controller vertical rate clearances (flight level and vertical rates), which reflects the human intervention. These clearances could be made available via the ATM system, Data Link or Voice Recognition.

Acknowledgments

The authors would like to thank Ferdinand Behrens from Technische Universität Berlin who managed the Full Flight Simulator sessions. Debora Fieberg from Technical University of Heidelberg developed the feature extraction software during her internal voluntary ship at DFS Deutsche Flugsicherung.

IEEE copyright notice: This article is a reprint and has been published in [25]. An article abstract is available in [26].

References

- [1] Christian Bousmanne et al., SESAR 2020 Concept of Operations Edition 2017, EUROCONTROL, November 2017.
- [2] Common Trajectory Prediction Capability for Decision Support Tools, Sip Swierstra, EUROCONTROL HQ, Brussels, Steven M. Green, National Aeronautics and Space Administration, Ames Research Center, Moffett Field, CA, 5th USA-Europe ATM Seminar, Budapest, June 2003.
- [3] A. Nuic. User Manual for the Base of Aircraft Data (BADA), Revision 3.12, EUROCONTROL, August 2014.
- [4] Ben Musialek, Carmen F. Munafo, Ryan Hollis, and Paglione Mike. Literature Survey of Trajectory Predictor Technology. Technical Report DOT/FAA/TCTN11/1, Federal Aviation Administration, William J. Hughes Technical Center, 2010.
- [5] Martin Lindner. Entwicklung eines Algorithmus zur Anpassung der Vorhersage von Trajektorien an spezifische Flugmanöver von Luftverkehrsgesellschaften. Technical University of Dresden, 2013.
- [6] Richard Alligier, David Gianazza, Nicolas Durand. Machine Learning applied to Airspeed Prediction during Climb. ATM Seminar 2015, 11th USA/EUROPE Air Traffic Management Seminar.
- [7] Richard Alligier, David Gianazza, Nicolas Durand. Ground based estimation of Aircraft Mass, adaptive vs. least square method. ATM Seminar 2013, 10th USA/EUROPE Air Traffic Management Seminar.
- [8] Areski Hadjaz, Gaetan Marceau, Pierre Saveant, Marc Schoenauer. Online Learning for Ground Trajectory Prediction. SESAR 2nd Innovation Days, 2012.
- [9] Yann Le Fablec and Jean Marc Alliot. Using Neural Networks to predict aircraft trajectories. In Proceedings of the International Conference on Artificial Intelligence (ICAI 99), Las Vegas, 1999.
- [10] Matthias Poppe et al. Flight Level Prediction with a Deep Feedforward Network. SESAR Innovation Days, Salzburg, December 2018.
- [11] Ian Goodfellow, Yoshua Bengio and Aaron Courville. Deep Learning. MIT Press Cambridge, Massachusetts, London, England, 2016.
- [12] DFS Deutsche Flugsicherung GmbH. Aeronautical Information Circular IFR 7, SSR Mode S Enhanced Surveillance. January 2003.
- [13] ICAO Doc 9871 AN/464 Technical Provisions for Mode S Services and Extended Squitter, First Edition, 2008.
- [14] Darren Rhodes & Spencer Norton. Aircraft Climb Performance. Presentation 2017.
- [15] Airbus Flight Operations. Getting to grips with Aircraft Performance. January 2002.
- [16] Debora Fieberg. Feature Preparation for a Machine Learning Prediction of Trajectories of Climbing Flights, Voluntary Internship at DFS, 2018.
- [17] Scikit-learn developers. Scikit-learn User Guide. Release 0.21.dev0, 2018.
- [18] Francois Collet et al. Keras, 2015. <https://keras.io>
- [19] Martin Abadi et al. TensorFlow: Large-scale machine learning on heterogeneous systems. 2015. Software available from tensorflow.org.
- [20] Trevor Hastie, Robert Tibshirani, Jerome Friedman. The Elements of Statistical Learning. Springer Series in Statistics, 2. Edition, 2008.

- [21] Sebastian Ruder. An overview of gradient descent optimization algorithms. Insight Centre for Data Analysis, NUI Galway. June 2017.
- [22] Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani. An Introduction to Statistical Learning. Springer New York Heidelberg Dordrecht London, 2013.
- [23] Rudolf Brockhaus, Wolfgang Alles and Robert Luckner, Flugregelung, Springer Heidelberg Dordrecht London New York, 2011.
- [24] Airbus Flight Operations. Getting to grips with cost index. Issue II - May 1998.
- [25] Matthias Poppe, Thomas Pütz & Roland Scharff. Improving Climb Performance Prediction in Air Traffic Control with Machine Learning and Full Flight Simulator Verification. In Proceedings of 38th Digital Avionics System Conference DASC. San Diego, USA, 2019.
- [26] Matthias Poppe, Thomas Pütz & Roland Scharff. Improving Climb Performance Prediction in Air Traffic Control with Machine Learning and Full Flight Simulator Verification. Abstract URL (in press): <https://ieeexplore.ieee.org>.

Tactical inputs for a stripless ATC system via Multi-Touch Gestures

Dr. Konrad Hagemann & Andreas Udovic

Introduction

„The DFS research centre is no ivory tower and doesn't want to be one. It is not about looking in the far future. The things that the colleagues are working on could already be implemented within the next 10 years.“ (cited after Ciupka, 2019). In accordance with this mission statement, DFS research centre focusses its research and development (R&D) efforts on technologies that are presumably mature enough to be deployed “on board” and to facilitate the ATCOs' work in the future.

There are big challenges that DFS will address within the next 10 years. One of them is the change to the stripless ATC system iCAS (iTEC Center Automation Systems) for lower airspace in German ATC centres, which will introduce new working procedures for the controllers. As part of its mission, the DFS research centre is investing significant effort in co-operations and programs on national and international level to develop concepts for future Air Traffic Control. In 2018, grant funding payments of €12.7 million were received for whole DFS from European research framework programmes, including SESAR and the German aeronautical research programme (DFS, 2019a, p. 20)

The Single European Sky ATM Research program (SESAR) is currently one of the most prominent Air Traffic Management (ATM) R&D initiatives worldwide (Tauss, 2017). The aim of SESAR 2020 project 16 is to reduce development and operating costs of Air Navigation Service Providers (ANSP). It achieves this by delivering the ATM Master Plan goals (defined by the European ATM community) for a more efficient deployment of human resources, a progressive increase in automation support, the implementation of virtualisation technologies and the use of standardised and interoperable systems whilst increasing the safety level. The PJ.16-04 “workstation, controller productivity” solutions deal with new methods of controller interaction with the Human Machine Interface (HMI), applying mature technologies from other domains to ATM. This will increase controller productivity, reduce workload, stress level and enable the use of SESAR advanced tools, safely facilitating performance based operations. One of these methods includes the use of Multi-Touch Inputs (MTI) at the controller working position (CWP, cf. SESAR, 2019a).

Background

Today, multi-touch technology is considered a mature technology, which is in everyday use in various environments (cf. SESAR, 2018). For example, tablets and smartphones are operated by multi-touch input and a standard set of gestures are known by most of the users (e.g. pinch to zoom). Products with multi-touch gesture interaction are often advertised with a promise to be faster, more direct and intuitive for the user. Nowadays, interaction via mouse, pen (paperless strip system, PSS) or single touch (voice communication system, SVS) are used for interaction with the ATC system at DFS (see also Hagemann & Udovic, 2014). So far, it has only been since the last couple of years that DFS put focus on the use of (multi-touch) gestures in the context of research and development activities. Several approaches have been taken to evaluate and pin-down the assumed benefits for ATCOs:

- In 2012, an early concept following a playful, multimodal approach has been realized as part of a student work (Wald, Buxbaum & Konopka, 2012).
- As part of a technology screening in SESAR 1, a direct and an indirect interaction concept involving different sets of multi-touch gestures have been researched (Hagemann & Udovic, 2014).
- Over a series of three validations DFS investigated an in-house built iCAS multi-touch Touch Input Display (TID) prototype that was capable of replacing the mouse completely (2-TID, c.f. Schilb, Perott & Schader, 2013a, 2013b, 2014).
- In 2015, DFS ordered a platform for the design of potential future ATCo workstations. The design-awarded Concept Desk can be operated with Multi-Touch gestures and in its first version even included the capability to process tangibles objects (Buxbaum, 2015)

Most recently, with SESAR solution 16-04 DFS developed a demonstration concept to integrate multi-touch gesture operations for frequent tactical inputs into a stripless system. The focus of this development was to:

- Demonstrate a multi-touch gesture interaction concept that may ease handling frequently required system inputs and that is ready for use
- Demonstrate together with INDRA its technical feasibility on the INDRA Industry-Based Platform (IBP).

Multi-Touch Interaction Concept

The Multi-Touch interaction concept has been developed over a series of six concept workshops together with ATCOs from different DFS ATC centres, software specialists and human factors experts as well as design experts from the University of Applied Sciences in Osnabrück. The interaction concept focussed on defining an intuitive and efficient solution for frequent tactical system inputs. Figure 11 illustrates the core elements of the concept by showing a typical workflow of selecting a label, opening a menu, selecting values and confirming it.

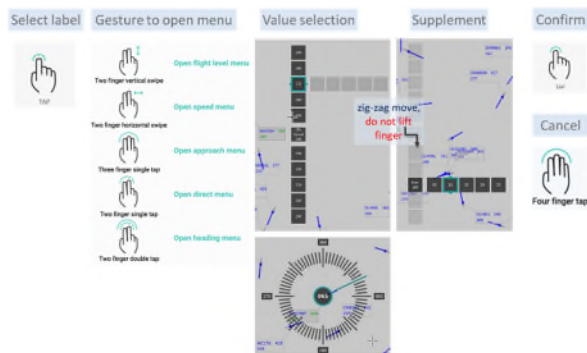


Figure 11: Typical workflow for multi-touch gestures.

The following clearances were possible via input from multi-touch gestures:

- Flight level with rate and less or greater
- Speed with less or greater
- Directs, i.e. proceed with direct course to a waypoint along the route
- Absolute Heading, i.e. heading in degrees
- Approach type, i.e. runway selected and approach type (VOR ILS, etc.)

In addition, functions for map handling (zoom in and out, centre map), label selection and interaction (shift label) as well as “assume” and “send” were solved by gestures. The complete set of gestures are presented in Table 2.

All multi-touch inputs could be done within a dedicated multi-touch area on the TID while the menus were shown on the radar display. To ease the learning of the gestures a help window could be manually triggered for display at the left side of the radar display. In addition,

Table 2: Complete set of gestures.

	1 Finger	2 Finger	3 Finger	4 Finger	5 Finger
Single Tap	(3) Select / Deselect Label	(5) "opening gesture": Display Direct menu	(8) "opening gesture": Display Approach Clearance menu	Cancel	
Double Tap	(4) Change Status Assume / Transfer etc."	"opening gesture": Display Heading menu	n.a. (It's error-prone)		
Swipe / Move	Move cursor on screen After opening gesture when menu for entry is displayed: select a value in menu	(6) "opening gesture": Level Menu: swipe up or down (7) "opening gesture": Speed Menu: swipe left or right	(9) after select: move label position	(1) Panning of map	
Press & Hold	(10) Mouse home	(11) Show trajectory"	Label reset: move selected label in default position	(1) Reset view (zoom, map position)	(2) Read-out
Pinch / thumb claw pinch	n.a.	(1) Zoom-in			
Reverse thumb claw pinch	n.a.	(1) Zoom-out			
Rotate	n.a.	After opening gesture: alternative entry of values			

some more complex inputs (e.g. combined clearances) were available and can displayed on the TID. However, they were rarely used in the technical demonstration as they would have required the ATCOs to look down from their radar screen.

The TID user interface was based on previous DFS prototype developments (Schilb, Perott & Schader, 2013a, 2013b, 2014) and has been updated to account for recent design guidelines. Overall, the complete layout of the menus in the mock-up and demonstrator was based on the “Styleguide for iCAS TID” developed by DFS (Tenoort & Hagemann, 2017).

Aim of the Demonstration

In sum, the following benefits were expected to be supported by the feedback collected during the demonstration:

The use of multi-touch gestures should...

- ...provide fast and efficient input for most frequent tactical inputs, i.e. clearances
- ...show high user acceptance through usability and attractiveness of the man-machine interface
- ...minimize head down time when using the TID for touch inputs together with the air situation window (ASW) for information display

The scope was to run a technical demonstration instead of a full-scale validation. Therefore, not all parts of the concept were implemented. Following this, a special challenge lied in realizing a show-case solution involving both a sufficiently complete concept mock-up and a running industrial prototype. In the end and due to constraints in development time, the latter could only show parts of the concept in operation on the IBP.

Methods

Participants

Thirty-one participants (4 female, 27 male) took part in the technical demonstration. The average age of the participants was 42 years with a job experience of 13.5 years. 17 ATCOs (11 Centre controllers, five approach controllers and one tower controller), six HMI experts and two software developers participated in the demonstration. In addition, eight other DFS employees took part in the event. The participants reported to have a very heterogeneous level of experience with using multi-touch. Five participants answered that they were not confident with using multi-touch gestures at all (overall mean = 3.29, SD = 1.76, see Figure 13, question 1)

Technical Set-Up

Two stations with prototypes representing different development steps (stand-alone concept vs. industry-based platform demonstrator) of almost completely the same user interface design concept (i.e. with a similar set of gestures) were used for the technical demonstration.

Figure 12 shows station 1, the SESAR concept demonstrator as jointly developed by DFS and the University of Applied Sciences in Osnabrück. The solid stand for ASW and TID has been especially designed for



Figure 12: Set-up of DFS mock-up during technical demonstration.

the purpose of this stand-alone technical concept demonstration. At this station a continuous radar replay was displayed, i.e. menus could be opened and values could be entered, but with no consequences on the flights.

For station 2 (see Figure 14), ATM system manufacturer INDRA had transferred the interaction concept on the SESAR IBP (iCAS System) together with a simulation capability, i.e. the input into the menus of the

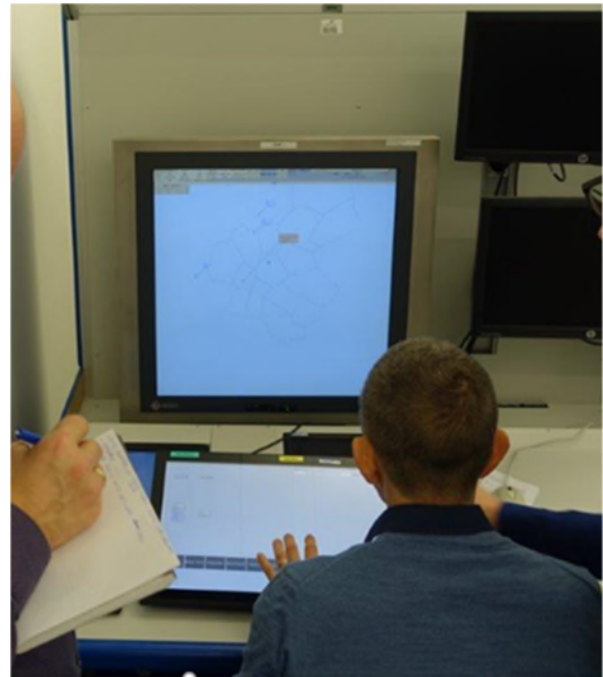


Figure 3: Set-up of INDRA iCAS IBP-based demonstrator during technical demonstration.

demonstrator were sent to the Flight Data Processing System (FDPS) so that a real behaviour of the system was realised. This demonstrates a basic technical feasibility of the developed solution in a real-life ATM System. At this station, the touch-monitor hardware was identical to the one that has been previously selected for the implementation of iCAS phase 2 at DFS.

The mock-up and the demonstrator were different in certain details: for example, the demonstrator did not provide an input via heading menu, instead only a picture was displayed. On the other hand, the demonstrator allowed the user to scroll up and down inside some of the ASW menus (e.g. flight level and speed menu), which was not implemented in the mock-up.

Schedule

Altogether ten demo slots of 1.5 h duration were run over five days.

Table 3: Schedule of 1.5h demo slot.

Time	Action	Responsibel
10min	Welcome and introduction	DFS
10min	Presentation of MTI concept	DFS/INDRA
<i>Splitting the participants in two groups</i>		
25min	Demonstrator INDRA with discussion and feedback Group 1	Demonstrator DFS with discussion and feedback Group 2
25min	Demonstrator DFS with discussion and feedback Group 1	Demonstrator INDRA with discussion and feedback Group 2
<i>Together as one group</i>		
20min	Final Questionnaire and debriefing	DFS/INDRA

Table 3 shows the schedule of one slot of the Technical Demonstration. It was designed such that participants were split-up in two groups who were testing the two prototypes in parallel. The briefing at the beginning and the debriefing at the end were done with the whole group. At the station, participants were encouraged to continuously speak out loud their impressions while observers took notes.

Results

Results from the questionnaire

Figure 13 to Figure 15 show the result of the questionnaire filled in by the participants. One central hypothesis for the demonstration was that the input via gestures is faster than by mouse and/or keyboard. The colleagues did not supported this assumption. The average of the answers related to this question was 3.34 (see question 5, in Figure 13). The participants rated the input speed to be slower than with mouse. Compared with the actual used Paperless Strip System (PSS), which is operated by a pen, the difference in speed is estimated even greater. Again, the speed was rated lower for input via gestures than for using the pen (average 3.91, see question 6 in Figure 13). In total, the input speed via gesture was rated with 2.97 in average as medium fast. Many of the participants with professional background as an ATCo are very well trained with using the pen for system input in their job. It is expected that with more training for the novel input modality the estimated input speed for multi-touch gestures will improve. Clearly, for a valid comparison of the different interaction modes a large scale validation exercise would be necessary.

Besides this, the overall concept was rated as suitable for handling an ATC system (average 2.31, question 2 in Figure 13).

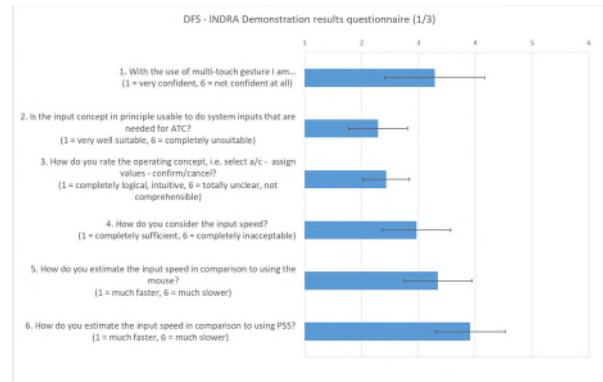


Figure 13: Results of DFS - INDRA Questionnaire (part 1)

Safety was another topic, which should be investigated in the demonstration. Therefore, two questions were asked to the participants. The first question addressed the error tolerance and the second question was related to the possibility to correct wrong inputs. The failure tolerance of the system was rated negatively by trend (mean = 3.70). The question related to the possibility to correct false input was estimated slightly more positive. A mean of 2.91 indicates that this point should be improved.

Another hypothesis for the demonstration was the assumption of a high usability of the gestures and menus. Figure 15 and Figure 16 show results related to this hypothesis. The participants rated all gestures and menus as suitable. Only the selection of the flight level in the flight level menu (question 11, see Figure 15) was rated in tendency more negative (mean = 2.97).

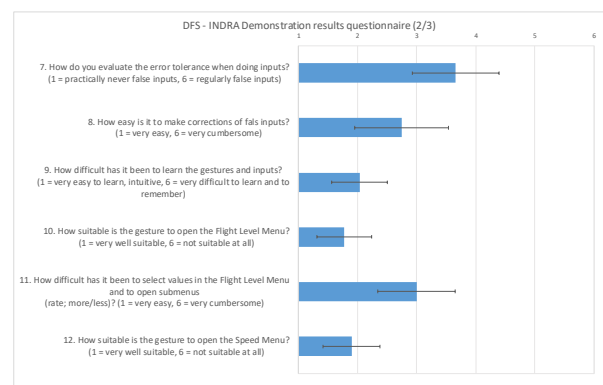


Figure 15: Results of DFS - INDRA Questionnaire (part 2)

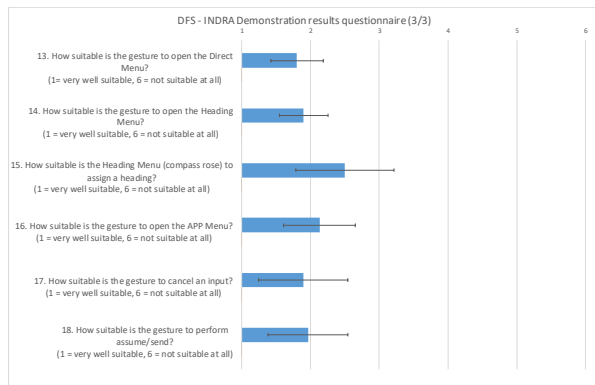


Figure 16: Results of DFS - INDRA Questionnaire (part 3)

Results from the verbal feedback during the demonstration and debriefing

All participants had the possibility to use the INDRA demonstrator and the DFS mock-up. Altogether, the participants provided many comments during the demonstration. During the interaction with the systems, the participants were encouraged to think aloud and give recommendations or ask questions. Colleagues from INDRA and DFS collected this feedback. At the end of the demonstration all participants have had the possibility to summarize their impressions in a debriefing. To structure the debriefing some questions were prepared to guide the discussion.

Overall, the participants were in favour of the concept. Also the presented HMIs (e.g. menus) were judged positive by the majority. The discussions revealed some additional ideas for specific aspects (e.g. scaling of heading menu, switch of gestures for certain functions, add another abort function in menus). The ideas were recorded for potential testing in future versions of the prototype. Furthermore ergonomic issues were commented like position of hand, strain by prolonged periods of interaction, haptic quality of the TID and hygienic. These aspects may need to be considered in more detail in the future.

Discussion

In summary, the multi-touch solutions developed by INDRA and DFS were successfully presented during the demonstration. Although at the beginning, the participants reported to have a very heterogeneous level of confidence with the use of multi-touch gesture input in general, the concept was rated as useful and shall be further investigated, e.g. in real-time simulations. This may speak for the intuitiveness of the approach. Nevertheless, details of the concept and HMI can be

optimized and improved (e.g. assignment of gestures to functions, heading inputs, touch screen surface). The participants gave a lot of comments and recommendations in this direction. The assumption that the inputs by multi-touch are faster than by mouse or pen was not supported by the participants. Therefore, the participants expressed their wish to test the multi-touch input concept in a validation exercise with long validation runs (two hours), including training and reference runs. In addition, possible effects of strain on the fingers and wrist can then be investigated. The failure tolerance of the concept was also criticized and has to be improved. A point that should also be investigated during a validation trial.

Outlook

Based on the collected data from three validation exercises plus one technical demonstration, the SESAR 2020 project partners, including leading European ATM industry, came to the conclusion that no more pre-deployment research is needed to implement multi-touch gestures at the CWP (SESAR, submitted). The technology itself is considered mature. Among others, DFS and INDRA successfully demonstrated the technical feasibility of a working multi-touch interaction concept on an industry-based prototype. As a consequence, within the SESAR project it is currently under discussion with the SJU to proceed directly from Technical Readiness Level (TRL) 4 for this solution, towards TRL 6 (i.e. "ready for deployment") by end of SESAR wave 1 in 2019.

On the other hand, for DFS that started R&D work on the topic in 2012, an implementation within a 10-year limit (see above Ciupka, 2019) is currently not in view. For the iCAS deployment within the next years, DFS focusses on using the mouse for tactical inputs and using single touch on the TID for strategic inputs. Although a technical feasibility could be demonstrated, jumping directly to deployment may still hold a high risk at the moment. From DFS point of view, the concept needs further validation in a highly realistic environment such as a full-scale real-time simulation including sufficient training and reference runs. As the technical demonstration has shown, the introduced multi-touch interaction concept for tactical inputs still holds some significant open gaps that could not yet be addressed and open questions such as physical strain over time can only be satisfactorily answered in real-time simulations.

More research and discussion about feasible innovation at the workstation is needed. Therefore, a permanent installation of the SESAR DFS demonstrator has been realized in the Advanced Function Simulator room of the DFS research centre in Langen. It is used to continue the discourse and to inform about the status of the research results at the end of SESAR 2020 wave 1 (e.g. for DFS future systems class). As an additional part of DFS dissemination activities, the demonstration concept has also been successfully transferred to another demonstration platform (2nd generation of concept desk, see Figure 17 and Buxbaum, 2015), which allows its presentation to a wider audience on events and to further discuss merits and flaws of the approach.



Figure 17: SESAR Multi-Touch Interaction concept successfully transferred and presented on the DFS Concept Desk Mini.

In the discussions about multi-touch gesture input a potential combination with Automatic Speech Recognition (ASR), has frequently been brought up in the debriefings. For example, voice input could help to automatically select targets on the radar screen, which would be quicker than using touch input and maybe even the mouse to move a cursor. Most recently, DFS ASR results on a target location assistance have shown a high reliability for the recognition of call signs spoken by the ATCo (SESAR, 2019b). To explore benefits of using multimodality in ATC, various studies have tried to shed more light on the discussion (e.g. TriControl Concept, Ohneiser, Jauer, Gürlük & Uebbing-Rumke, 2016; Eurocontrol, 2007). Research in this area should be continued to get a better understanding and quantify the possible benefits and limitations for the ATCos. A claim that is supported by the European ATM Master Plan, which explicitly lists R&D activities to see “the development of new human machine interface (HMI)

interaction modes in relation to other SESAR Solutions (including new user interface technologies, such as speech recognition, multi-touch and gaze detection)” (SESAR, 2015, p.52).

In sum, much potential lies in the further optimization of the ATCo workplace and in the provision of a modern human-machine interaction concept for ATC systems that may meet the high expectations of current and upcoming controller generations. A challenge that the DFS research centre intends to address in more depth in future R&D activities and projects.

Abbreviations

ANSP	Air Navigation Service Providers
ASR	Automatic Speech Recognition
ASW	Air Situation Window
ATCo	Air Traffic Controller
ATM	Air Traffic Management
ATS	Air Traffic Services
CWP	Controller Working Position
FDPS	Flight Data Processing System
HMI	Human Machine Interface
IBP	Industrial Based Platform
iCAS	iTEC Center Automation Systems
MTI	Multi-Touch Inputs
PSS	Paperless Strip System
R&D	Research and Development
SESAR	Single European Sky ATM Research program
SVS	Sprachvermittlungssystem
TID	Touch Input Display

References

- Buxbaum, J. (2015). Der „concept desk“ im DFS-Forschungszentrum – ein Lotsenarbeitsplatz „von übermorgen“ zum Anfassen. *Innovation im Fokus*, 02/15, Langen: Deutsche Flugsicherung.
- Ciupka, S. (2019). Weder Elfenbeinturm noch Glaskugel. *Direct*, Ausgabe 1, Langen: Deutsche Flugsicherung.

- DFS (2019a). *Annual Report 2018*. Langen: Deutsche Flugsicherung. URL retrieved on May 14, 2019: https://www.dfs.de/dfs_homepage/en/Press/Publications/DFS%20Annual%20Report%202018_EN.pdf
- EUROCONTROL (2007). Multimodal interfaces: a brief literature review. *EEC Note No. 01/07*.
- Schilb, K., Perott, A. & Schader, N. (2013a). iCAS Validierung Bremen/ Unterer Luftraum Zwischenbericht 1. Validierung. Version 1.0. Bremen: Deutsche Flugsicherung.
- Schilb, K., Perott, A. & Schader, N. (2013b). iCAS Validierung Bremen/ Unterer Luftraum Zwischenbericht Intermediate Simulation. Version 1.0. Bremen: Deutsche Flugsicherung.
- Schilb, K., Perott, A. & Schader, N. (2014). iCAS Validierung Bremen/ Unterer Luftraum Abschlussbericht Final Simulation. Version 1.0. Bremen: Deutsche Flugsicherung.
- Hagemann, K. & Udovic, A. (2014). Bewertung von Multi-Touch Gesteneingaben am Lotsenarbeitsplatz. *Innovation im Fokus*, 01/14, Langen: Deutsche Flugsicherung.
- Ohneiser, O. Jauer, M-L, Gürlük, H. & Uebbing-Rumke, M. (2016). TriControl – Multimodal Air Traffic Controller Working Position. *The Sixth SESAR Innovation Days*, 08.-10. Nov. 2016, Delft, Netherlands.
- SESAR (2014). Innovation Analysis Report 2013. Deliverable 10.10.02-D93.
- SESAR (2015). European ATM Master Plan. The Roadmap for Delivering High Performing Aviation for Europe. Executive View. URL retrieved on May 14, 2019: <https://ec.europa.eu/transport/sites/transport/files/modes/air/sesar/doc/eu-atm-master-plan-2015.pdf>
- SESAR (2018). SESAR 2020 PJ.16-04-01 TRL2 TVALR – Multi-Touch Inputs (MTI). Deliverable D3.1.020.
- SESAR (submitted). SESAR 2020 PJ.16-04-01 TRL4 TVALR – Multi-Touch Inputs (MTI). Deliverable D3.2.020.
- SESAR (2019a). Controller working position/Human machine interface - CWP/HMI. URL retrieved on May 14, 2019: <https://www.sesarju.eu/projects/cwphmi>.
- SESAR (2019b). SESAR 2020 PJ.16.04.02 TRL4 TVALR – Automatic Speech Recognition (ASR). Deliverable D3.2.020. Appendix A.
- Tauss, G. (2017). Die F&E-Programme SESAR 1 und SESAR 2020 – Rückblick und Ausblick. *Innovation im Fokus*, 02/17, Langen: Deutsche Flugsicherung.
- Tenoort, S. & Hagemann, K. (2017). Styleguide for iCAS TID - SFIT. *Innovation im Fokus*, 02/17, Langen: Deutsche Flugsicherung.
- Uebbing-Rumke, M., Gürlük, H., Jauer, M.-L., Hagemann, K. & Udovic, A. (2014). Usability Evaluation of Multi-Touch-Displays for TMA Controller Working Positions. *Fourth SESAR Innovation Days 2014*, 25.-27. Nov. 2014, Madrid, Spain.
- Wald, D., Buxbaum, J. & Konopka, J. (2012). Innovative Mensch-Maschine Interaktionsmöglichkeiten in der Flugverkehrskontrolle. *TE im Fokus*, 1/12, Langen: Deutsche Flugsicherung.

Acknowledgments

This article reports about data from a project, which has received funding from the SESAR Joint Undertaking under the European Union's Horizon 2020 research and innovation programme under grant agreement No 734141.

Disclaimer

This article reflects only the authors' view. The SESAR JU is not responsible for any use that may be made of the information it contains.

Erkenntnisse aus dem Projekt “StayCentered - Methodenbasis eines Assistenzsystems für Centerlotsen“

Jörg Buxbaum

Einleitung

Monitoringsysteme zur Zustandsmessung von Menschen haben in den letzten Jahren in sprachbasierten Dialogsystemen und Entertainmentsystemen Einzug gehalten [4, S. 15]. Im Bereich Automotive gelten diese Systeme oftmals schon als Voraussetzung für einen teilautonomen Betrieb eines Fahrzeugs [9]. Ein weiteres mögliches Anwendungsfeld von Emotionserkennung liegt in der Arbeit von Fluglotsen der Streckenkontrolle. Im Vordergrund steht hier die Fragestellung, ob sich der Operateur in der Überlast befindet und z.B. von umliegenden Sektoren entlastet werden sollte. Günstigstenfalls ist ein derartiges System in der Lage, Überlast vorherzusagen, damit rechtzeitig Gegenmaßnahmen getroffen werden können.

Im BMBF-Verbundprojekt „StayCentered – Methodenbasis eines Assistenzsystems für Centerlotsen“ hat die Technische Universität Chemnitz Ansätze für eine Emotionserkennung bei Fluglotsen entwickelt und auf ihre Einsatzmöglichkeiten zur Zustandserfassung hin untersucht. In dem Projekt (Laufzeit: 2015 – 2018) hatte die DFS die Rolle des Industriepartners inne und stellte Lotsen, Simulationsinfrastruktur und passende Verkehrsszenarien für Versuche zur Verfügung.

Im Folgenden werden die Erkenntnisse aus den Projektberichten zusammengefasst, eine individuelle Bewertung der Versuchsansätze vorgenommen sowie ein Resümee aus Sicht der Flugsicherung gezogen.

Monitoring und bedarfsweise Unterstützung der Fluglotsen aktuell

Bereits heute wird von Fluglotsen in Kontrollzentralen der Stresslevel von direkten Kollegen sowie benachbarter Sektoren akustisch sowie über Gesten und Posen wahrgenommen [1]. Diese Wahrnehmung des „Arbeitszustands“ auch von Kollegen in Nachbarsektoren folgt dem Verständnis der Lotsen, dass sich der Teamgedanke nicht nur auf die Dyade Executive-Planner bezieht, sondern auch auf die Kooperation der Lotsen innerhalb der Einsatzberechtigungsgruppe (EBG) und darüber hinaus [5]. Bei entsprechender Indikation verfügen die Kollegen benachbarter Sektoren auch heute schon über wirksame Möglichkeiten, die Workload

des überlasteten Sektors zu verringern: z.B. durch Übernahme von Verkehr oder Umleitungen um den betroffenen Sektor herum. Es bestehen derzeit ebenfalls schon technische Möglichkeiten, um Hoch- oder Überlast anderen Kollegen im Betriebsraum zu signalisieren. Dazu dient z.B. im Center München die so genannte „Genfer Ampel“ – eine Eingabemöglichkeit am Lotsenarbeitsplatz, die anderen Sektoren eine Lastinformation anzeigen kann. Allerdings wird die „Genfer Ampel“ bei hoher Arbeitslast der Lotsen nicht genutzt, da Lotsen gerade unter Hochlast auf vermeidbare zusätzliche Systemeingaben verzichten [1]. Das System wird also gerade dann nicht verwendet, wenn es eine Wirkung erzielen soll.

In DFS-Entwicklungssimulationen zur Untersuchung neuer Systeme, Verfahren und Lufträume wird die Workload der Lotsen im Regelfall über NASA TLX / ISA durch den Lotsen selber erfasst. Die Verwendung des gebräuchlichen ISA-Fragebogens beeinflusst allerdings die Arbeit der Fluglotsen und kann Effekten von Fehleinschätzungen oder Fehlverhalten unterworfen sein [2]. In anderen Simulationen (AcListant, DLR, 2016 [19]) wurden Lotsen Zusatzaufgaben gegeben, deren Zeitbedarf für die Erledigung als Indikator für verbleibende kognitive Reserven genutzt wurden. Eine Erfassung der Workload auf Basis einer psychophysiologischen Zustandsmessung findet u.a. aufgrund des damit verbundenen Aufwands in Entwicklungssimulationen bei DFS und DLR nur sehr selten statt.

Versuche

Versuchsreihen und Datenerhebungen zur Lotenzustandsmessung wurden in drei Phasen im DFS-Forschungszentrum vorgenommen: November 2015, September 2017 und Dezember 2017 [4, S. 15]. Zum Einsatz kam die CATO-Simulationsumgebung (Controller Assistance Tools), die an das Air-Traffic-Management-System (ATM-System) P1 bzw. P2 der DFS angelehnt ist. Dabei kamen u.a. Hochlastverkehrsszenarien zum Einsatz, die eine hohe Workload bei den Fluglotsen forcieren sollten.

Folgende Messungen wurden vorgenommen:

1. Videoaufzeichnung der Gesichter der Fluglotsen per HD-Webcam zur Erkennung mimischer

Emotionsausdrücke per FACS (Facial Action Coding System) [18]

2. Videoaufzeichnung der Fluglotsen per stereoskopischer Kamera für die weitere Auswertung von Posen und Gesten
3. Videoaufzeichnung des kompletten Lotsenarbeitsplatzes per HD-Webcam zur Identifizierung von Interaktionen zwischen den Lotsen
4. Messung des Blickbewungsverhaltens inklusiver Erfassung der Pupillendilatation per Blickbewegungsbrille zur weiteren Auswertung der kognitiven Belastung sowie der Interaktion mit dem Arbeitsplatz
5. Messung des zeitlichen Verlaufs der Hautleitfähigkeit und der Herzrate zur weiteren Diagnostik des kognitiven und emotionalen Erlebens
6. Audioaufzeichnung der verbalen Kommunikation der Center-Lotsen zur weiteren Diagnostik des kognitiven und emotionalen Erlebens
7. Erfassung der subjektiven Belastung (ISA) und des subjektiven Situationsbewusstseins im zeitlichen Verlauf

Weitere Messungen fanden u.a. in einem experimentellen Arbeitsumfeld an der TU Chemnitz statt. Die dortigen Probanden waren Studenten der Universität ohne dezidierte berufliche Vorkenntnisse oder Kenntnisse aus der Praxis der Flugverkehrskontrolle.

Ergebnisse

Gestik / Posen

Beobachtungen während Simulationen sowie in Arbeitssituationen in der Kontrollzentrale München zeigten deutliche Indikatoren für die herrschende Arbeitslast des Betriebspersonals. Im realen Betrieb führen Phasen mit geringem Verkehrsaufkommen zu einer entspannten, zurückgelehnten Haltung. Körperhaltung von Planner und Executive unterschieden sich dabei aufgrund der überwiegenden Blickrichtung (Planner: Streifenlage, Executive: Air Situation Display) und der eingesetzten Arbeitsmittel. So legt der Executive aufgrund der Verwendung des Mikrofons den Arm oft auf der Konsole auf und arbeitet deswegen tendenziell nach vorn gebeugt – unabhängig von der Belastungssituation [4, S. 20]. Bei hohem Stresslevel sitzen Fluglotsen eher aufrecht [1]. Lotsen können aufgrund der individuellen Körperhaltung einzelner Kollegen erkennen, ob diese sich in einer Stresssituation befinden und können darauf

reagieren [4, S. 20]. Folgerichtig ist dies bei fehlendem Sichtkontakt zum Kollegen nicht möglich.

Auf Basis der in Simulationen erfassten Daten gelang es nicht, für alle Lotsen eine Korrelation zwischen Arbeitssituation und gemessenen, spezifischen körpersprachlichen Signalen zu ermitteln. Ergebnisse von körpersprachlichen Analysen ließen sich von Experimenten mit Laien / Studenten dazu nicht auf Fluglotsen übertragen [4, S. 20]. Im Vergleich zu Laien ist die Gestik von Fluglotsen während der Arbeit spärlich ausgeprägt. In Versuchen in der Kontrollzentrale München nahmen 65% der Fluglotsen über die Versuchsdauer von rund einer Stunde hinweg lediglich eine einzige Pose ein [4, S. 21]. Die Auswertung der Körpersprache als Indikator für Stress fällt demnach als Möglichkeit aus [4, S. 21]. Beitragender Faktor dafür war auch der Einfluss der unterschiedlichen Beleuchtung und Sitzposition des Lotsen auf die Qualität der erfassten Bilder, was zu einer geringen Fehlertoleranz des Algorithmus führte.

Mimik (Facial Action Coding System, FACS)

Die Analyse der Mimik über das Facial Action Coding System ergab für übliche Situationen bei der Fluglotsenarbeit keine nutzbaren Ergebnisse und wurde deswegen nicht in das Emotionsmodell integriert [4, S. 57]. Beitragender Faktor für die „sparsame Mimik“ könnte sein, dass Centerlotsen ausgewählt und trainiert sind, professionell und ruhig mit emotional herausfordernden Situationen umzugehen. Eine emotionale Erregung könnte daher deutlich schwieriger zu messen sein als es bei Standardprobanden der Fall ist [4, S. 14]. In diesem Sinn könnte auch ein hoher Grad an gewissenhafter Ausübung einer beruflichen Tätigkeit ein möglicher Indikator für einen hohen Grad an emotionaler Selbstkontrolle sein. Tatsächlich zeigte sich dies an mehreren Stellen im Projekt. So wiesen beispielsweise im Vergleich von Experimenten mit Studenten und Fluglotsen die Fluglotsen eine signifikant höhere Ausprägung der Persönlichkeitseigenschaft Gewissenhaftigkeit auf. Im Vergleich zu den Studenten zeigten sie geringere Veränderungen in der positiven und negativen Emotion [4, S. 70].

Audio

Die durchgeführte Analyse des Sprachsignals zur automatischen Erkennung von Emotionen führte zu keinen unmittelbar verwendbaren Ergebnissen. Potentiell nutzbar scheint die Emotionserkennung über die Analyse der sprachlichen Semantik, die allerdings den Einsatz einer zuverlässigen bilingualen Spracherkennung

(englisch und deutsch) am Lotsenarbeitsplatz voraussetzt Eine derartige, zuverlässige Spracherkennung ist derzeit nicht verfügbar. Allerdings wäre dabei zu berücksichtigen, dass zum einen private Gespräche zwischen Lotsen nicht analysiert werden, zum anderen findet ein Teil der Kommunikation über funktionale Gesten statt, die logischerweise bei der Spracherkennung unberücksichtigt bleiben [4, S. 40, S. 57, S. 59]. Die vielfach standardisierte Kommunikation von Fluglotsen könnte jedoch grundsätzlich unterstützen, per Spracherkennung untypische Äußerungen zu ermitteln und daraus ggf. besondere Situationen zu erkennen [4, S. 14].

Pupille und Hautleitfähigkeit

Als potentiell relevanteste psychophysiologische Parameter erwiesen sich die Hautleitfähigkeit sowie die Pupillenerweiterung. Die Pupillengröße erwies sich bei schweren gegenüber leichten Aufgaben als signifikant vergrößert [4, S. 43, S. 60] [3].

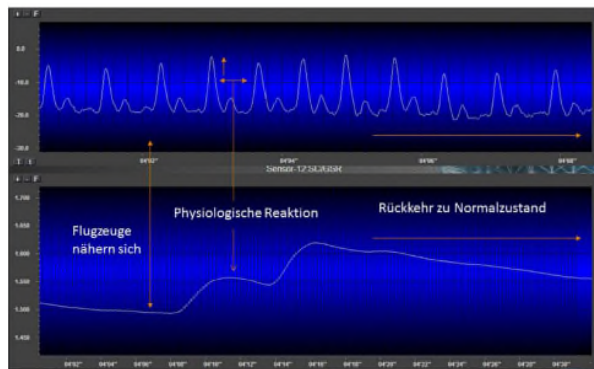


Abbildung 4: Prototypische Zuordnung eines Events zum psychophysiologischen Erregungspotentialen am Beispiel der Pupillendilatation. [s. 4, S. 44, Abb. 17]

Für die Hautleitfähigkeit konnten schon in den Pretests bekannte Effekte erfolgreich repliziert werden, d.h. mit Zunahme der Beanspruchung zeigte sich aufgrund der erhöhten Schweißsekretion eine unmittelbare Steigerung der Hautleitfähigkeit [4, S. 44, Abb. 16].

Schlussfolgerungen

Auf Basis der Ergebnisse des Projektes ergeben sich Schlussfolgerungen mit Relevanz für die weitere Forschungsarbeit, die im folgenden Teil des Artikels dargestellt werden.

Resümee zu den Methoden

Von den eingesetzten Messmethoden erwiesen sich allein die Pupillendilatation und der Wert der Hautleitfähigkeit als geeignete Indikatoren für

Arbeitslast. Alle anderen Methoden lieferten in der gewählten Experimentalkonstellation unbrauchbare Resultate, trotz teilweise aufwändiger Datenauswertung, teilweise unter Zuhilfenahme von Maschinellern / KI. FACS und Audio erkannten in vielen Fällen sogar jeweils unterschiedliche Emotionen [4, S. 57].

Ethische, rechtliche und soziale Implikationen (ELSI)

Mit Blick auf die Nutzerakzeptanz wurden in einem ELSI-Workshop im November 2016 in der Kontrollzentrale Bremen die Zustandsmessungen an sich sowie die Verwendung der erhobenen Daten diskutiert. Generell waren die Lotsen mit den Messungen einverstanden, wenn das Ziel die Verbesserung der Arbeitsbedingungen ist – beispielsweise im Rahmen von Untersuchungen neuer Arbeitsverfahren oder Assistenzsysteme. Dabei sollte die Einhaltung des Datenschutzes hohe Priorität haben. Abgelehnt wurde die Verwendung der Daten im Betrieb – hier sehen Fluglotsen das Risiko des Datenmissbrauchs und die Verwendung zu anderen Zwecken als der unmittelbaren Unterstützung bei der Flugverkehrskontrolle [4, S. 74] [2].

In anderen Gesprächen äußerten Lotsen, dass die Information über den aktuellen, gemessenen, persönlichen Stresswert für den Lotsen selber hilfreich sein könne um z.B. einzuschätzen, welche zusätzliche Aufgabe (z.B. Annahme von Sonderanfragen durch Lfz-Besatzungen) oder Verkehrsmengen man zu leisten in der Lage sei und welche man ablehnen solle [4, S. 75]. Supervisoren als eine der Zielgruppen der Nutzerzustandsinformation formulierten hingegen in Interviews, dass sie keine Emotionsdaten von Lotsen in Echtzeit angezeigt bekommen möchten. Grund dafür sei zum einen der Schutz der Privatsphäre der Lotsen – und auch die erkannte Schwierigkeit bei der korrekten Interpretation der Daten, sowie dem damit verbundenen Risiko, dass die Informationen zu unangemessenen Entscheidungen führen könnten. Verwendet werden solle lediglich der Arousalwert als Indikator der Arbeitsfähigkeit des Lotsen generell. Starkes Interesse bestünde bei den Supervisoren an der unmittelbaren Erkennung von Extremsituationen wie z.B. extremen Stress in einzelnen Sektoren [4, S. 82].

Bei Überlegungen zum langfristigen operativen Einsatz einer Lotsenzustandsmessung wird eine Abwägung zwischen einer potentiell verbesserten Schadensvermeidung und deutlichen Eingriffen in Persönlichkeitsrechte nötig sein [2]. Zu einem gewissen Grad finden solche Eingriffe in die Persönlichkeitsrechte bereits heute im Rahmen von Alkohol- und Drogentests für

operative Mitarbeiter in verantwortungsvollen Arbeitsrollen statt (Piloten, Zugführer), ohne dass darüber ein gesellschaftlicher Disput herrscht. Der frühzeitige Einbezug der Arbeitnehmervvertretung in potentielle Entwicklungen wird allerdings genauso dringend empfohlen, wie eine transparente Durchführung einer etwaigen Forschung, Entwicklung und Implementierung. Es bietet sich die Chance, bei potentieller Anwendung derartiger Messungen im Simulationsumfeld diese Kultur bereits für alle beteiligten Parteien erlebbar zu machen. In diesem Verständnis war zum ELSI-Workshop in Bremen auch der örtliche Betriebsrat eingeladen worden.

Erkenntnisse zum experimentellen Design

Mangels der Möglichkeit, Experimente im betrieblichen Umfeld durchzuführen, wurden zum einen wie beschrieben Experimente an einem Simulator der DFS durchgeführt. Um weitergehende Untersuchungen mit anderen Schwerpunkten vorzunehmen, erstellten die Projektteilnehmer der TU Chemnitz vereinfachte, teilweise transportable Experimentalumgebungen, die Lotsenarbeitsplätzen nachempfunden waren, allerdings auf einem hohen Abstraktionslevel und mit einem hohen Maß an Vereinfachung. Im Rahmen von Experimenten an diesen Lotsenarbeitsplatz-Mockups ergab sich eine wichtige Erkenntnis:

Experimentelle Designs mit möglichen Kollisionen von Flugzeugen sind nicht sinnvoll. Zum einen handelt es sich um eine Situation, die bei Fluglotsen extrem negativ belegt ist und eine große mentale Belastung auslöst. Besser wäre ein experimentelles Design, bei dem schlimmstenfalls ein vertikales, automatisches Ausweichmanöver mit Notstaffelung simuliert wird. Zum anderen zeigte sich in Experimenten mit 25 Studenten der Universität Pompeu Fabra (Barcelona), dass bei Probanden, die nicht als Fluglotse ausgebildet waren, nach einer Mid Air Collision makaberer Weise die Workload sank, da ja plötzlich zwei Flugzeuge weniger zu kontrollieren waren [3].

Adaptive Anpassung des Nutzerinterfaces und generelle Anforderungen an das HMI

Gemäß Erkenntnissen des Projekts würde während fordernder Verkehrssituationen eine Anpassung des Interfaces den Lotsen unterstützen, Überlast zu vermeiden. Das CWP-Display könne dazu an den Lotsenzustand angepasst werden (adaptives Interface), beispielsweise bei hohem Stress oder Unterforderung [1]. Im Projekt wurden dabei insbesondere der

Vereinfachung und angemessene Reduktion der Datendarstellung in Stresssituationen eine wichtige Rolle zugeschrieben [4, S. 77, S. 80]. Eine automatische Anpassung des Interfaces unter Hochlastbedingungen sahen die Fluglotsen jedoch kritisch [4, S. 80].

Die Forscher identifizierten ebenfalls Maßnahmen, um belastenden Situationen einfacher vorbeugen zu können:

- Schnellere Informationsaufnahme der hoch-relevanten Höheninformation im Luftlagebild mittels geeigneter Darstellung (derzeit: rein numerische Darstellung)
- Darstellung von zu erwartenden Einflügen, die sich noch außerhalb des Radarbildausschnitts befinden
- Generelle Überarbeitung des Arbeitsplatzdesigns (u.a. Reduktion Anzahl der Monitore)
- Bedarfsgerechte Distanzinformationen zwischen ausgewählten Flügen auf dem Luftlagebild
- Übersichtlichere Wetterdarstellung

Nach [1] sind für den Streckenlotsen Informationen in folgender Rangfolge für seine Arbeit wichtig: Höhe, Steigrate/Sinkrate, laterale Position, Heading, Ground Speed. Im User Interface sollte sich diese Reihenfolge widerspiegeln – was derzeit gerade in Bezug auf die Höhe nicht erkennbar ist, die auf dem Bildschirm rein numerisch dargestellt ist. Ob sich annähernde Flugzeuge ggf. im Begriff sind, Staffelungswerte zu unterschreiten, erkennt der Fluglotse derzeit, indem er die etwa 5 mm großen Ziffern auf dem Air Situation Display abliest und miteinander vergleicht.

Aufbauend auf Beobachtungen und Interviews insbesondere mit Fluglotsen der Kontrollzentralen Langen, München und Karlsruhe (Februar – August 2015) ergab sich für die Forscher die Perspektive, dass das derzeitige operationelle Arbeitsumfeld deutliche Verbesserungspotentiale birgt. So äußerten Fluglotsen im Rahmen der Untersuchungen mehrfach ihren Unmut über die Erreichbarkeit von Informationen und Funktionalitäten und insbesondere über die allgemeine Übersichtlichkeit der aktuellen Nutzerschnittstellen, was gemäß der Forscher der TU Chemnitz eine generelle Überarbeitung der Darstellungsweise und Interaktionen nahelegt [4, S. 75]. Auch an anderen Stellen zeigte sich, dass die Arbeitsweise der Fluglotsen und die Ausgestaltung des ATM-Systems nicht ausreichend aufeinander abgestimmt sind. So werde gemäß Lotsenangaben (in Beobachtungen und Interviews August 2015) im System VAFORIT bzw. iCAS Karlsruhe der Bildschirm mit den Flugstreifen-Informationen „nicht genutzt“. Bei Beobachtungen im Betriebsraum im Rahmen von „StayCentered“ stellte sich allerdings

heraus, dass die Lotsen den Bildschirm dennoch nutzten [5]. Schließlich berge die leicht erfassbare Anzahl der „Streifen“ die Möglichkeit, die in den nächsten 10 bis 20 Minuten zu erwartende Arbeitsbelastung abzuschätzen. Auch wurden die Streifen nach Höhe sortiert, um die vertikale Auslastung des Sektors zu illustrieren. Damit zeigt sich erneut die in [1] dargelegte Relevanz der Darstellung der Höheninformation für Streckenlotsen.

Eine weitere Herausforderung ergab sich gemäß den Projektergebnissen dadurch, dass Fluglotsen der Streckenkontrolle beim System P2 meist nicht die Luftlagedarstellung (Air Situation Display, ASD) zur Erkennung potentieller Konflikte nutzten [1]. Das ASD wird vorwiegend verwendet, um zu erkennen ob sich Flugzeuge wie erwartet verhalten. Die Konflikterkennung findet auf Basis einer internen Repräsentation des Verkehrsbildes („Picture“) statt, die rund 3 Minuten in die Zukunft weist. Demzufolge müsste die Arbeitsposition der Lotsen so gestaltet werden, dass Executive und Planner dieses „Picture“ unmittelbar teilen können. Allerdings hängt dieses „Picture“ von individuellen, persönlichen mentalen Modellen ab, so dass eine Übertragung in ein technisches System eine Herausforderung darstelle [1]. Vergleichbar ist diese Aufgabe damit, bei einem laufenden Schachspiel nicht nur die Positionen der Figuren darzustellen, sondern die von den Spielern in Erwägung gezogenen Züge (= Freigaben oder erwartete bzw. herbeigeführte Trajektorien) bzw. das (wahrscheinlichste) Spielbild nach zwei oder drei Zügen.

Marktsituation und vergleichbare Forschungsvorhaben

Mit zunehmender Automatisierung sowie immer besseren Möglichkeiten, mittels leistungsstarken Rechnern Sensordaten in Echtzeit analysieren zu können, hat die Zustandsmessung auch in anderen Anwendungsdomänen in den letzten Jahren deutliche Fortschritte erzielt.

So hat Medienberichten zufolge die Firma Harman International ein kamerabasiertes Augen- und Pupillenverfolgungssystem entwickelt, das die Pupillenerweiterung als Indikator der mentalen Belastung eines Autofahrers misst [7]. Das neue proprietäre Augen- und Pupillenverfolgungssystem soll hohe kognitive Belastung und mentales Multitasking im Fahrersitz messen können, und den anderen Sicherheitssystemen des Autos signalisieren, sich dem Zustand des Fahrers anzupassen. Eine auf den Fahrer gerichtete Kamera soll komplexe Sensoren in Sitzen und Lenkrädern, oder biometrischer Sensoren, die physischen Kontakt mit

dem Fahrer benötigen, überflüssig machen. Die Kamera erfasse kontinuierlich die Pupillenerweiterung des Fahrers und ein proprietärer Software-Algorithmus analysiere den Pupillenreflex mit Hilfe einer adäquaten Filter- und Signalverarbeitung. Wie die Entwickler berichten, isoliere und identifiziere der Filter durch hohe kognitive Belastung ausgelöste Reaktionen. Die berechneten Ergebnisse sollen demnach verwendet werden können, um intuitiv die Einstellungen der Benutzer-Interfaces anzupassen, um sowohl physische als auch psychische Ablenkung des Fahrers zu minimieren, etwa durch Versetzen mobiler Geräte in den "Do not disturb"-Modus oder die Anpassung der Interventionsschwelle des Adaptiven Fahrerassistenzsystems (ADAS) [7].

Im laufenden Forschungsvorhaben „F 2402-II Validierung einer kontinuierlichen Methode zur Beanspruchungs-ermittlung am Beispiel des Fluglotsenarbeitsplatzes“ der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA) und dem DLR setzen die Forscher zur Zustandsmessung auf gellose EEG-Systeme in Verbindung mit maschinellem Lernen [10]. Ergebnisse werden im Rahmen der Abschlussveranstaltung des Projekts Ende Januar 2020 erwartet.

Das Start-Up „Tawny AI“ [11] aus München konzentriert sich bei der Entwicklung einer Zustandsmessung von Menschen auf rein optische Methoden, verbunden mit Maschinellern Lernen. Nach Aussage der Geschäftsführung im Rahmen einer Präsentation auf dem Serviceplan-Innovationstag am 25.9.2019 reiche für eine Emotionserkennung die Analyse von Videomaterial. Dabei stellt Tawny AI heraus, dass für Anwendungen in sicherheitskritischen Aufgaben (Pilot, Fluglotse, Betriebspersonal Kraftwerk) weniger Emotionen wie Trauer, Wut oder Freude relevant seien, sondern Unter- und Überforderung sowie „Flow“, dem als beglückend empfundenen Gefühl des völligen, konzentrierten Vertiefens in eine Tätigkeit [12]. Tawny AI berichtet, dass gerade im Flow die Ablenkung des Lotsen z.B. durch Arbeitslasterfassung per Eingabe in ein System (z.B. NASA TLX) negativ auf den Flowzustand wirke und demzufolge beitragen könne, dass positiv empfundener Stress in negativ empfundenen Stress umschlägt – verbunden mit einer rapiden Abnahme der Arbeitsleistung und der Aufmerksamkeit. Diese Sicht deckt sich mit den gewonnenen Erkenntnissen aus StayCentered.

Bewertung und Ausblick

Im Projekt „StayCentered“ wurde ein zeitgemäßes Methodenrepertoire praxisnah daraufhin untersucht, ob

eine Messung des Beanspruchungszustands von Fluglotsen der Streckenkontrolle technisch möglich ist und ggf. auch eine darauf aufbauende Prognose von Beanspruchungszuständen ermöglicht wird.

Als Ergebnis schieden die meisten Ansätze aus. Teilweise lag dies wohl auch daran, dass Lotsen im Rahmen ihrer betrieblichen Aufgabenerfüllung signifikant weniger Emotionalität oder auffällige Mimik / Gestik erkennen lassen. Lediglich die Pupillendilatation und der Wert der Hautleitfähigkeit erwiesen sich als geeignete Indikatoren für Arbeitslast. Eine zuverlässige Prognose der Arbeitslast konnte nicht umgesetzt werden.

Als Konsequenz aus den Untersuchungen bietet sich an, Systeme, die auf Pupillendilatation und Hautleitfähigkeit basieren, im Rahmen von Simulationen im DFS-Forschungszentrum testweise zur Erfassung der Arbeitslast einzusetzen. Voraussetzung dafür ist ein Screening der technischen Angebote auf dem Anbietermarkt. Speziell auch im Anwenderbereich „Automotive“ wäre eine Recherche sinnvoll, wo im Rahmen der aktuell dort stattfindenden weitgehenden Automatisierung die größten Entwicklungssprünge im Bereich „Sensorik für Zustandsmessung von Operateuren“ in naher Zukunft zu erwarten sind.

Als Herausforderung einer sensorgestützten Beanspruchungsmessung zeigt sich, dass schwankende kognitive und emotionale Befindlichkeiten der Lotsen Auswirkungen auf „tagesformabhängige“ Anstrengung und Stressempfinden haben [4, S 58]. Je nach Befindlichkeit unterscheidet sich demnach die subjektive Einschätzung einer objektiv identischen Arbeitslast, was im Emotionsmodell als individueller Faktor zu berücksichtigen wäre. Wie eine tagesgenaue Festlegung dieses Faktors für jeden einzelnen Fluglotsen gestaltet werden könnte müsste noch geklärt werden. Allerdings ist dieser Umstand ja auch schon jetzt bei der „händischen“ Erfassung von empfundener Arbeitslast relevant und erschwert die exakte Bewertung von Arbeitslast in heutigen Experimenten beispielsweise an Simulatoren der Streckenkontrolle.

Weitere Erkenntnisse, die eher am Rande des Projekts gesammelt wurden, weisen auf massive Entwicklungspotentiale bei der zukünftigen Ausgestaltung der Nutzerschnittstelle von Fluglotsen hin. Ebenso formulierten die Forscher der TU Chemnitz die Annahme, dass eine adaptive und an die Arbeitslastsituation angepasste Nutzerschnittstelle helfen könne, kurzfristige Überlastsituationen sicher zu meistern – ähnlich wie es die Entwicklungen im Bereich Automotive bereits indizieren [7].

Abkürzungsverzeichnis

ADAS	Adaptive Driver Assistance System
ASD	Air Situation Display
ATM	Air Traffic Management
BMBF	Bundesministerium für Bildung und Forschung
CATO	Controller Assistance Tools
CWP	Controller Working Position
EEG	Elektroenzephalogramm
FACS	Facial Action Coding System
HMI	Human Machine Interface
iCAS	iTEC Center Automation Systems
ISA	Instantaneous Self Assessment
KI	Künstliche Intelligenz
Lfz	Luftfahrzeug
NASA TLX	NASA Task Load Index
VAFORIT	Very Advanced Flight Data Processing Operational Requirement Implementation

Referenzen

- [6] Pfeiffer, L., Valtin, G., Müller, N. H. & Rosenthal, P. (2015). Aircraft in Your Head: How Air Traffic Controllers Mentally Organize Air Traffic.
- [7] Truschzinski, M., Pfeiffer, L. & Valtin, G., (2017). ELSI-Aspekte in Forschungsverbünden. In: Burghardt, M., Wimmer, R., Wolff, C. & Womser-Hacker, C. (Hrsg.), Mensch und Computer 2017 - Workshopband. Regensburg: Gesellschaft für Informatik e.V.. DOI: 10.18420/muc2017-ws07-0378
- [8] Truschzinski, M. (2017). Modellierung und Vorhersage von mentaler Arbeitsbeanspruchung in einem Fluglotsenaufgabenexperiment, Informatik 2017, Gesellschaft für Informatik, Bonn.
- [9] Brunett, G., Eibl, M., Hamker, F., Ohler, P., Protzel, P. (2018). Schlussbericht StayCentered – Methodenbasis eines Assistenzsystems für Centerlotsen (MACeLot), Technische Universität Chemnitz.
- [10] Pfeiffer, L., Müller, N. H. (2016). The Mental Organization of Air Traffic and its Implication to an

- Emotion Sensitive Assistance System, International Journal on Advances in Life Sciences, 8 (1 & 2): 164-174.
- [11] Pfeiffer, L., Müller, N. H., Valtin, G., Truschzinski, M., Ohler, P. & Rosenthal, P. (2015). Emotionsmodell für zukünftige Mensch-Technik-Schnittstellen zur Unterstützung von Centerlotsen. DGLR-Bericht 2015-01.
- [12] Hofacker, A. (2016). Pupillen-basiertes System misst kognitive Belastung des Fahrers, URL aufgerufen am 12.11.2019: <https://www.springer-professional.de/fahrerassistenz/pupillen-basiertes-system-misst-kognitive-belastung-des-fahrers/7821044>
- [13] Blöcher, T. (2018), „Der prüfende Blick, Kamera-basierte Fahrerzustandserkennung“, , URL aufgerufen am 12.11.2019: <https://www.elektro-niknet.de/elektronik-automotive/assistentensysteme/der-pruefende-blick-157495-Seite-2.html>
- [14] ABLresearch (2016). Camera-Based Driver Monitoring Systems to be Chief Enablers of Safe, Semi-Autonomous Driving, URL aufgerufen am 12.11.2019: <https://www.abiresearch.com/press/camera-based-driver-monitoring-systems-be-chief-en/>
- [15] BAuA: Bundesanstalt für Arbeitsschutz und Arbeitsmedizin. Experimentelle Untersuchungen zur Entwicklung der kontinuierlichen neuronalen Beanspruchungsmessung für den Feldeinsatz Projektbeschreibung, Projektnummer F 2402, Teilprojekt F 2402-II „Validierung einer kontinuierlichen Methode zur Beanspruchungsermittlung am Beispiel des Fluglotsenarbeitsplatzes“, URL aufgerufen am 12.11.2019: <https://www.baua.de/DE/Aufgaben/Forschung/Forschungsprojekte/f2402.html>
- [16] Website der TAWNY GmbH, URL aufgerufen am 12.11.2019: <https://tawny.ai/>
- [17] Csíkszentmihályi, M. (1995). Flow. Das Geheimnis des Glücks. 4. Auflage. Klett-Cotta, Stuttgart.
- [18] Ekman, P. & Friesen, W. V. (1978). Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, Palo Alto, CA
- [19] Projektwebseite AcListant (Active Listening Assistant), URL aufgerufen am 27.11.2019: <http://www.aclistant.de/wp>

Innovation im Fokus Informationen zu Forschung, Entwicklung und Validierung

Innovation im Fokus erscheint mindestens jährlich und beschäftigt sich bevorzugt mit Informationen zu Forschung, Entwicklung und Validierung mit Beteiligung DFS Deutsche Flugsicherung GmbH. Diese Ausgabe ist elektronisch im Internet (www.dfs.de > [Flugsicherung](#) > [F&E](#) > [Service](#)) sowie über das DFS Intranet verfügbar. 90 Exemplare werden in gedruckter Form aufgelegt.

Wie auch der Vorgänger-Zeitschrift „TE im Fokus“ wurde dieser Zeitschrift von der Deutschen Bibliothek eine ISSN (International Standard Serial Number) zugeteilt:

Printversion: ISSN 2198-8951 (vormals: 1861-6364)

Internet-Version: ISSN 2198-896X (vormals: 1861-6372)

Datum dieser Ausgabe: 10.12.2019

DISCLAIMER

Alle hier erwähnten Warenzeichen sind Eigentum ihrer jeweiligen Besitzer. Warenzeichen werden nicht ausdrücklich als solche gekennzeichnet. Aus dem Fehlen von Urheber- oder Markenrechtskennzeichen darf jedoch nicht geschlossen werden, dass es sich um einen nicht geschützten Namen oder um eine nicht geschützte Marke handelt.

COPYRIGHT

Dieses Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwendung außerhalb der Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung der DFS Deutsche Flugsicherung GmbH unzulässig und wird zivil- und strafrechtlich verfolgt. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Copyright © 2019 by DFS Deutsche Flugsicherung GmbH - Weitergabe sowie Vervielfältigung dieser Unterlage, Verwertung und Mitteilung ihres Inhalts nicht gestattet, soweit nicht ausdrücklich zugestanden. Zuwiderhandlungen verpflichten zu Schadensersatz. Alle Rechte vorbehalten, insbesondere für den Fall der Patenterteilung oder GM-Eintrag.



Impressum

Innovation im Fokus

Informationen zu Forschung,
Entwicklung und Validierung

Herausgeber:

DFS Deutsche Flugsicherung GmbH

Günter Achatz, Bereichsleitung
Planung & Innovation;

Redaktion:

Dr. Konrad Hagemann
Tel. +49 (0)6103 707 5745
E-Mail: konrad.hagemann@dfs.de

Stefan Tenoort
Tel. +49 (0)6103 707 5769
E-Mail: stefan.tenoort@dfs.de

Dr. Morten Grandt
Tel. +49 (0)6103 707 1139
E-Mail: morten.grandt@dfs.de

Oliver Haßa
Tel. +49 (0)6103 707 5762
E-Mail: oliver.hassa@dfs.de

Anschrift der Redaktion:

DFS Deutsche Flugsicherung GmbH
Redaktion Innovation im Fokus
Am DFS-Campus 5
63225 Langen
E-Mail: forschung@dfs.de

Nachdruck nur mit Genehmigung.

