

2/11

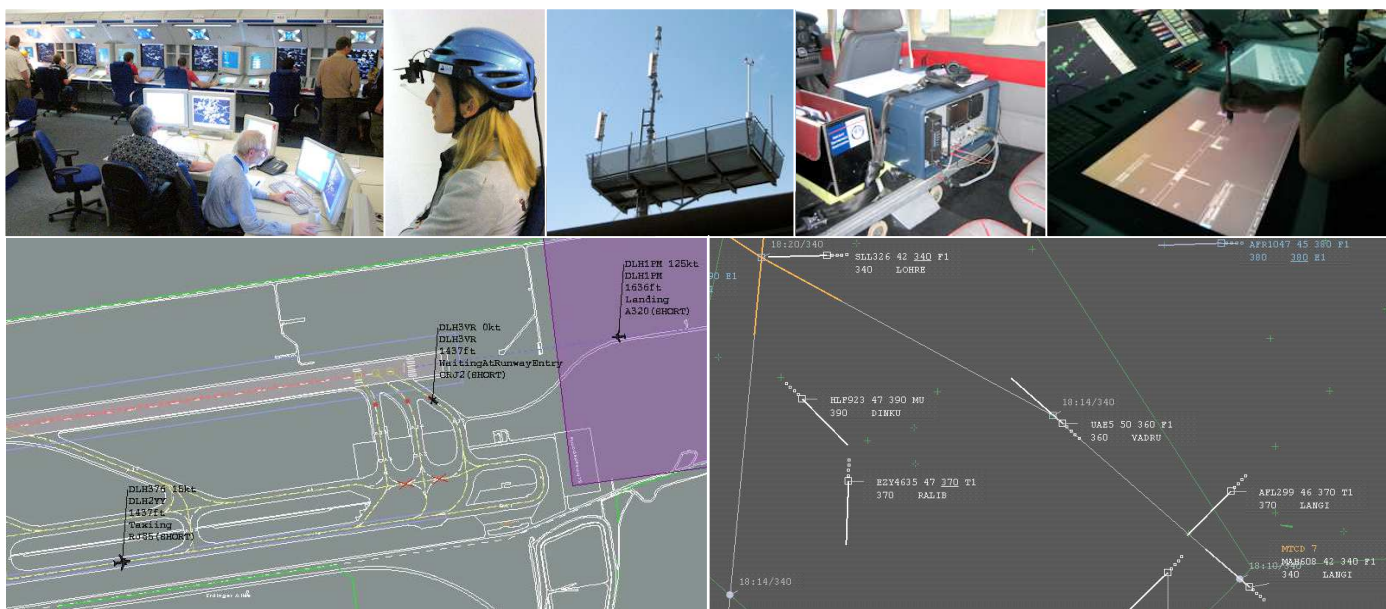
06.12.2011

TE *im Fokus*

Informationen aus dem Bereich Forschung und Entwicklung der DFS Deutsche Flugsicherung GmbH

In dieser Ausgabe:

- Vergleich der Blickerfassungsqualität dreier Eyetrackingsysteme
- Kopplung zwischen einem Schnellzeitsimulationstool und betrieblichen Planungssystemen
- Kalibrierung und Validierung eines taskbasierten Arbeitslastmodells für Flugsicherungssysteme der DFS



DFS Deutsche Flugsicherung

Inhalt

Abstracts der Heftbeiträge	2
---	----------

Jianhua Lambrecht, Dr. Konrad Hagemann, Andreas Udovic, Dr. Morten Grandt Vergleich der Blickerfassungsqualität dreier Eyetrackingsysteme	3
---	----------

Katharina Schwentek Kopplung zwischen einem Schnellzeitsimulationstool und betrieblichen Planungssystemen	15
---	-----------

Dr. Ralph Leemüller, Stefan Tenoort, Theresa Emsbach Kalibrierung und Validierung eines taskbasierten Arbeitslastmodells für Flugsicherungssysteme der DFS	26
--	-----------

Impressum.....	41
-----------------------	-----------



Die DFS Deutsche Flugsicherung GmbH
ist nach DIN EN ISO 9001:2008 zertifiziert.

Abstracts der Heftbeiträge

Jianhua Lambrecht, Dr. Konrad Hagemann, Andreas Udovic, Dr. Morten Grandt: **Vergleich der Blickerfassungsqualität dreier Eyetrackingsysteme** - Dieser Bericht stellt die Ergebnisse aus einer vergleichenden Untersuchung dreier Eyetrackingsysteme hinsichtlich Messgenauigkeit und Datenverluste dar. Ziel der Untersuchung war es, zu prüfen, inwieweit aktuell auf dem Markt erhältliche, kontaktlose Eyetracking-systeme einen Mehrwert für Untersuchungen der Mensch-Maschine Interaktion (MMI) im Flugsicherungskontext liefern können. Das Ergebnis zeigt, dass die Blickbewegung als Eingabemedium aufgrund messtechnisch bedingter Einschränkungen im Bereich der Flugverkehrskontrolle (FVK) nicht einsetzbar ist. Jedoch werden andere potentielle Anwendungsfelder z.B. zur Kapazitätsanalyse, zur Analyse von Blickverhaltensmustern, zur allgemeinen Systemgestaltung sowie zur Erfassung des mentalen Zustandes des Nutzers praxisnah vorgestellt und diskutiert.

Katharina Schwentek: **Kopplung zwischen einem Schnellzeitsimulationstool und betrieblichen Planungssystemen** - Die optimale Ausnutzung verfügbarer Kapazitäten an Hub-Flughäfen spielt mit dem wachsenden Luftverkehrsaufkommen eine immer bedeutendere Rolle. Bei bisherigen Untersuchungen im Rahmen nationaler Forschungsprogramme lag der Fokus auf der Weiterentwicklung von einzelnen taktischen (AMAN, DMAN) und prätaktischen Planungssystemen (CLOU/FMAN) für Hub-Flughäfen. Im Verbundvorhaben iPort sollen diese Planungssysteme sinnvoll miteinander gekoppelt werden, um den potentiellen Nutzen einer prätaktischen Mengenvorsteuerung für den Verkehrsfluss zu untersuchen. Als Werkzeug zur Durchführung einer automatisierten Simulation dient das Schnellzeitsimulationstool AirTop. Ziel ist ein Vergleich der Verkehrsabwicklung in Hinblick auf Durchsatz und Delay ohne und mit einer CLOU/FMAN Mengenvorsteuerung, welche die Sequenzplanung der taktischen Systeme beeinflusst. Im Rahmen dieses Projekts wird zum ersten Mal ein Schnellzeitsimulationstool mit (Echtzeit-) Planungssystemen gekoppelt.

Dr. Ralph Leemüller, Stefan Tenoort, Theresa Emsbach: **Kalibrierung und Validierung eines taskbasierten Arbeitslastmodells für Flugsicherungssysteme der DFS** - Das bisher genutzte Arbeitslastmodell für Flugsicherungssysteme der DFS wurde durch ein taskbasiertes Arbeitslastmodell zur Ermittlung von Lotsen-Belastungen und Sektor-Kapazitäten ersetzt. Der vorliegende Beitrag beschreibt das entwickelte Workload-Modell. Darüber hinaus wird der Prozess der Validierung der Workload-Parameter für die Niederlassung Karlsruhe mit der seit 2011 zur Verfügung stehenden P1/VAFORIT-Umgebung dargestellt. Das taskbasierte Workload-Modell erbringt auf der Basis der neuen Task-Parameter plausible Ergebnisse für die Kriterien Belastung und Kapazität und bildet die Belastung durch das genutzte Flugsicherungs- (Air Traffic Control-, ATC-)System adäquat ab. Eine Verschiebung der Beanspruchung innerhalb des Teams vom Radar- zum Planungsloten wurde anhand der Kalibrierung bestätigt. Das entwickelte Modell ist eine Grundlage für die Berechnung von Kapazitätswerten für die Sektoren der Niederlassung Karlsruhe mit dem System CAPAN (Capacity Analyser).

Vergleich der Blickerfassungsqualität dreier Eyetrackingsysteme

Jianhua Lambrecht (TUD), Dr. Konrad Hagemann, Andreas Udovic, Dr. Morten Grandt (DFS)

Zusammenfassung

Dieser Bericht stellt die Ergebnisse aus einer vergleichenden Untersuchung dreier Eyetrackingsysteme hinsichtlich Messgenauigkeit und Datenverlustrate dar. Ziel der Untersuchung war es, zu prüfen, inwieweit aktuell auf dem Markt erhältliche, kontaktlose Eyetrackingsysteme einen Mehrwert für Untersuchungen der Mensch-Maschine Interaktion (MMI) im Flugsicherungskontext liefern können. Das Ergebnis zeigt, dass die Blickbewegung als Eingabemedium aufgrund messtechnisch bedingter Einschränkungen im Bereich der Flugverkehrskontrolle (FVK) nicht einsetzbar ist. Jedoch werden andere potentielle Anwendungsfelder z.B. zur Kapazitätsanalyse, zur Analyse von Blickverhaltensmustern, zur allgemeinen Systemgestaltung sowie zur Erfassung des mentalen Zustandes des Nutzers praxisnah vorgestellt und diskutiert.

1. Einleitung

Trotz vielversprechender Ideen für Anwendungsfelder von Eyetrackingsystemen, z.B. als Eingabemedium spielen eine hohe Messgüte und eine geringe Datenverlustrate eine grundlegende Rolle in der Entscheidung für oder gegen den Ein-

satz dieser Technologie.

Die Anwendung von Eyetrackingsystemen verbindet die Forschungsgebiete der Physiologie, der Psychologie bis hin zur Mensch-Maschine-Kommunikation und der Informatik. Der Einsatz des Eyetrackings im Bereich der Bildschirmarbeit ermöglicht es, Aussagen darüber zu treffen, an welcher Bildschirmposition welche Informationen aufgenommen werden und wie viel Zeit die Informationsaufnahme in Anspruch nimmt. Außerdem ist es möglich, verschiedene Strategien der Informationssuche zu unterscheiden. In ähnlicher Weise können vigilanz- oder beanspruchungs-korrelierte psychophysiologische Effekte aufgedeckt werden, z.B. anhand der Fixationsfrequenz und -dauer oder mittels okulomotorischer Parameter wie Lidschlussdauer und -frequenz. Weitergehende Ansätze beschäftigen sich mit der Idee, Blickbewegungen als ein erweitertes Eingabemedium zu nutzen. Aus einem 2009 bei der DFS durchgeführten Ideenworkshop stammen die in Abbildung 1 dargestellten Anwendungsfelder von Blickbewegungsdaten im Bereich der Flugverkehrskontrolle. Im Folgenden werden in Ergänzung hierzu relevante Ergebnisse der Blickbewegungsforschung ausführlicher dargestellt.

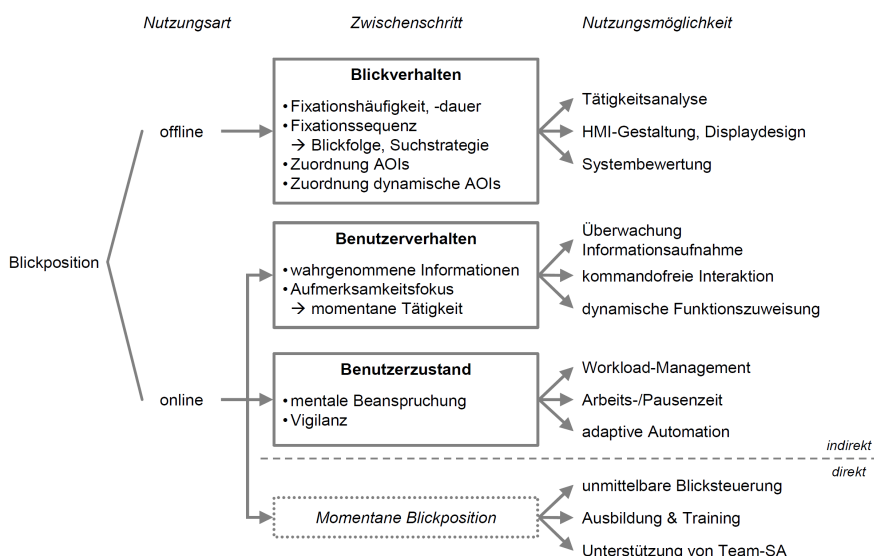


Abb.1: Nutzungsmöglichkeiten von Blickbewegung im FVK-Bereich (Grandt et al., 2010).

1.1 Blickbewegung als Eingabegerät zur Kontrolle von Benutzungsschnittstellen

Viele Entwicklungsarbeiten sind von dem Gedanken getragen Blicksteuerung als Eingabegerät zu verwenden, um Menschen zu unterstützen, die z.B. aufgrund körperlicher Einschränkungen keine Maus oder Tastatur bedienen können (Jacob & Karn, 2003; Bach-Y-Rita & Kercel, 2003).

Allgemeine Vorteile des Einsatzes von Blickbewegungen als Eingabegerät hat Drewes (2010) zusammenfassend dargestellt:

- Die Steuerung mit den Augen ist schneller als die Steuerung über andere Eingabegeräte. Eine Kombination von den Standard-Computereingabegeräten Maus und Tastatur mit einer Blicksteuerung ermöglicht eine schnelle Interaktion zwischen Mensch und Maschine. Sibert und Jacob (2000) haben zwei Experimente durchgeführt, um den Unterschied zwischen Blick- und Maussteuerung eingehender zu untersuchen. Dabei wurden den Teilnehmern Such- und Selektionsaufgaben gestellt und die beiden Eingabemodalitäten miteinander verglichen. Die Ergebnisse zeigten, dass diese Aufgaben mit der Blicksteuerung schneller erledigt werden konnten.
- Die Blicksteuerung hat eine höhere räumliche und zeitliche Flexibilität, da die Augen immer verfügbar sind. Insbesondere in einer Umgebung, in der die Hände für manuelle Eingaben nicht immer unmittelbar verfügbar sind. Drewes erwähnt hierzu ein Beispiel aus der medizinischen Anwendungsforschung. So hat ein Chirurg am Operationstisch beide Hände aus Hygienegründen oftmals nicht verfügbar, und über Blicksteuerung bestünde die Möglichkeit zu einer erweiterten, berührungslosen Interaktion.

Zum Einsatz von Eyetracking als Eingabemedium am Bildschirmarbeitsplatz sind bisher nur wenige Arbeiten veröffentlicht worden.

2006 wurde am NLR eine Studie durchgeführt (Esser & Bos, 2006a, 2006b), in der Blickfixationen dazu verwendet wurden, um Luftfahrzeuge auf einem Radardisplay vorzuselektieren und anschließend Geschwindigkeit, Flughöhe oder Heading auszuwählen. Für 12 Probanden konnte gezeigt werden, dass eine Vorselektion mittels Blicken signifikant schneller und fehlerfreier

durchführbar war als die herkömmliche Selektion mittels Trackball. Dieser Effekt war direkt proportional zur Distanz der zu selektierenden Luftfahrzeuge untereinander, was den Ansatz insbesondere für große oder mehrere verteilte Bildschirme vielversprechend erscheinen lässt. Auch die Aussagen der Probanden deuteten darauf hin, dass diese Methode gegenüber der herkömmlichen Selektionsmethode allgemein bevorzugt wird.

Eine Eingabe von Flugplandaten mittels Blicksteuerung erwies sich jedoch gegenüber der herkömmlichen Methode als weniger effizient. Die Gründe hierfür vermuten die Autoren in der Art und Weise der Implementierung, die als stark invasiv¹ wahrgenommen wurde, in der zu geringen Stabilität des Datenoutputs des Eyetrackers und nicht zuletzt darin, dass Blickeingaben für diese spezielle Art der Aufgabe womöglich generell schlecht geeignet sind. Leider sind die Ergebnisse der Studie nur eingeschränkt auf den FVK-Bereich übertragbar: So wurden lediglich 19"-Monitore eingesetzt und ein kopfbasiertes Eyetrackingsystem verwendet; solche kopfbasierten Systemen verfügen bekanntermaßen über eine höhere Genauigkeit als kontaktlose Systeme.

Im Kontext der Flugverkehrskontrolle sind den Autoren keine spezifischen Arbeiten mit kontaktlosem Eyetracker und einem 2K*2K Standardarbeitsmonitor bekannt. Es erschien daher naheliegend, eine kleine Demonstration Ende 2009 in Rahmen einer Realzeitsimulation mit Lotsen aus der DFS-Niederlassung München durchzuführen, um einen ersten Eindruck über bestehende Möglichkeiten dieser Idee zu gewinnen. Dabei wurde prototypisch ein Blickbewegungsmessgerät der Firma Tobii mit einem En-Route Radar Simulator gekoppelt. In dieser Demonstration konnten Blickbewegungen des Lotsen für die Steuerung des Mauszeigers auf dem Radarbildschirm verwendet werden. Das Fazit der sich anschließenden Diskussion mit den Fluglotsen lautete, dass die Blicksteuerung als Ersatz von Computermouse oder Tastatur in der dargestellten Form wenig geeignet erschien (Grandt et al., 2010). Das Hauptproblem lag darin begründet, dass die gezeigte Implementierung als hochinvasiv empfunden wurde. Vom Nutzer wurde dies als stark ab-

¹ beeinträchtigend, hinderlich, störend

lenkend empfunden, da der Cursor den Augen ständig folgte und schwierig zu positionieren war. Das Anvisieren eines Ziels durch Blicksteuerung und ein anschließendes akkurates Klicken per Maus waren nur schwer möglich. Die Ergebnisse zeigen die zentrale Bedeutung einer ausreichenden Genauigkeit bei der Erfassung der Messdaten durch das Gerät und das Risiko einer zu hohen Invasivität bei der Implementierung.

Die allgemeinen Probleme mit der Nutzung von Blicksteuerung als Eingabegerät wurden bereits in einer Reihe von Studien diskutiert. Herkömmliche Eingabegeräte (z.B. Maus) stellen mindestens zwei Funktionen zur Verfügung: ein Kommando auswählen und dieses z.B. mittels Mausklick initiieren. Mit dem Auge lässt sich hingegen nichts unmittelbar „anklicken“. Wenn hingegen durch das einfache Betrachten des Bildschirms ständig Aktionen ausgelöst werden, entsteht das sog. „Midas-Touch-Problem“ (Jacob, 1990; Jauch & Karn, 2003), bei dem der Benutzer unabhängig davon, wo er hinblickt, immer eine Aktion auszulöst. Die einfachste Möglichkeit zur Vermeidung dieses Problems ist die Verwendung von Latenzzeiten, d.h. das Auslösen zu verzögern und abubrechen, sobald der Benutzer seinen Blick abwendet. Der Einfluss auf eine mögliche schnellere Ermüdung der Augen ist dabei allerdings noch eingehender zu untersuchen.

1.2 Blickbewegung als Indikator für den mentalen Zustand

1.2.1 Pupillendurchmesser und mentale Beanspruchung

Mentale Beanspruchung bezieht sich vor allem auf die menschliche Informationsverarbeitung, d.h. kognitive Prozesse wie Wahrnehmung, Rechnen, Sprache und Gedächtnis (Kahnemann, 1973). Der Grad der mentalen Beanspruchung hängt einerseits von der Komplexität der Aufgabe, andererseits von den individuellen Leistungsvoraussetzungen ab.

Zur Erfassung der mentalen Beanspruchung werden häufig Fragebögen oder psychophysiologische Methoden verwendet. Mit Fragebögen kann die subjektive Bewertung der Beanspruchung erfasst werden, während psychophysiologische Methoden die Art und Stärke mentaler Beanspruchungszustände mittels psychophy-

siologischer Variablen objektiv messen. Die physiologische Veränderung (z.B. kardiovaskuläre Reaktionen, elektrodermale Aktivität, Augenbewegungsaktivität) reflektiert die Beanspruchung d.h. die körperliche Reaktion auf die Belastung. Aus der Literatur sind verschiedene Blickbewegungsindikatoren für mentale Beanspruchung bekannt. Im Folgenden wird insbesondere genauer auf den Pupillendurchmesser eingegangen, da hierzu bereits eine größere Zahl an Untersuchungen vorliegen.

Laut Bösel (1987) reguliert die Pupillenweite drei Faktoren der visuellen Informationsaufnahme: die Beleuchtungsichte im Auge, die Abbildungsschärfe (Schärfentiefe) und den Abbildungs- kontrast auf der Netzhaut. Der Zusammenhang zwischen Pupillenbewegung und mentaler Beanspruchung ist in einigen Studien bestätigt worden (z.B. Beatty, 1982; Steinhauer et al., 2004). Mit zunehmender Schwierigkeit der Aufgabe weitet sich die Pupille (Hyönä et al., 1995; Just & Carpenter, 1993). Aber es gibt auch kritische Stimmen. So kritisierte Schultheis (2004) die Allgemeingültigkeit der bisherigen Untersuchungsergebnisse, da diese streng laborkontrolliert waren und nur gleichartige und kurze Aufgaben enthielten, über die dann bei der Auswertung gemittelt wurde. Er überprüfte die empirisch bestätigten Ergebnisse und versuchte sie auf alltägliche Anwendung zu generalisieren. In seinem Experiment sollten die Probanden verschiedene Texte (zwei leichte und zwei schwierige Texte) nacheinander lesen und nach jedem Text einige Verständnisfragen beantworten. Er kam aufgrund seiner Untersuchung zu der Schlussfolgerung, dass kein stabiler Zusammenhang zwischen Pupillengröße und mentaler Beanspruchung bei der Verarbeitung längerer Texte nachweisbar war. Demnach wäre die Pupillengröße als Indikator für eine Online-Erfassung mentaler Beanspruchung nicht geeignet. Wie der Autor selbst einräumt, bleibt jedoch zu klären, inwieweit methodische Mängel der Untersuchung Einfluss auf die Ergebnisse hatten oder ob Pupillengröße und mentale Beanspruchung zumindest im Bereich der Sprachverarbeitung lediglich nur über eine sehr kurze Zeit korrelieren.

Beim pupillometrischen Verfahren sind einige weitere Einflussfaktoren bekannt, die sich auf die Messergebnisse auswirken können. So zählt Krüger (2000) über 20 Faktoren auf, die sich auf die Pupille auswirken. Einer der stärksten Ein-

flussfaktoren auf die Pupillenweite ist die herrschende Leuchtdichte. Dabei beträgt die auf Grund der Änderung der Leuchtdichte maximal hervorgerufene Änderung des Pupillendurchmessers zirka 8 mm (Galley, 2001), während die Änderung der mentalen Beanspruchung eine Änderung des Pupillendurchmessers von nur zirka 2 mm hervorruft. Zudem können auch individuelle Faktoren Störvariablen für die Messung darstellen. Dazu gehören das Alter (Birren, 1950; Kumnik, 1954), die Einnahme von Medikamenten oder Alkohol, Krankheit (Gambill et al., 1967), Pigmentverschiebung der Hell-Dunkeladaptation, Nah-Akkommodation (Hornung, 1967), Farbeinfluss und körperliche Anstrengung (Bouma, 1962) sowie das subjektive Lichtempfinden (Galley, 2001).

Aufgrund der Ergebnisse vorhergehender Untersuchungen erscheint das pupillometrische Verfahren zum aktuellen Kenntnisstand als nicht geeignet, mentale Beanspruchung am Lotsenarbeitsplatz zu ermitteln. Darüber hinaus bleiben noch viele Fragen zu klären, wie z. B. unter welchen Bedingungen und wie schnell die Pupillen auf einen Reiz reagieren. Einerseits beeinflussen wie oben erwähnt, vielfältige Störvariablen aus der Umwelt und dem Individuum die Messqualität. Andererseits ist der visuelle Kanal bei der Flugkontrollarbeit stark belastet. Aufgrund dessen ist schwer zu unterscheiden, ob die Pupillen auf eine hohe Anzahl visueller Reize oder auf eine kognitive Belastung reagieren.

1.2.2 Lidschlag, Vigilanz und visuelle Beanspruchung²

Unter dem Begriff „Vigilanz“ versteht man Zustände der Daueraufmerksamkeit. Eine Reaktion auf zufällige, schwelennahe, selten auftretende

Ereignisse setzt Vigilanz voraus. Zu den negativen Beanspruchungsfolgen zählt die Ermüdung, die unmittelbar zu einer Reduktion der Vigilanz führt. Schleicher et al. (2008) untersuchten den Zusammenhang zwischen Augenaktivität und Schläfrigkeit beim Fahren. Die Parameter Lidschlagdauer, Verzögerung der Wiederöffnung des Lidschlags, Lidschlagintervall sowie standardisierte Geschwindigkeit des Lidschlusses wurden als die besten Indikatoren für sowohl subjektiv berichtete als auch objektiv gemessene Schläfrigkeit identifiziert. Außerdem zeigten Parameter der Blicksprünge (Sakkaden) und Fixationsdauern spezifische Veränderungen mit zunehmender Schläfrigkeit.

Des Weiteren führte Weber (2002) eine Untersuchung mit 20 Fluglotsen durch, um aktuelle und zukünftige Flugsicherungssysteme (u.a. der DFS) hinsichtlich Ermüdung und visueller Beanspruchung von Fluglotsen zu vergleichen. In seiner Untersuchung konnte er einen Zusammenhang zwischen verringerter Lidschlagrate und der visuellen Beanspruchung darlegen. Auch führte Ermüdung bei lang andauernden Aufgaben oder durch Fehlbeanspruchungen zu einem Anstieg der Lidschlagrate.

1.3 Einsatz von Blickbewegungen in der Validierung

Neben den oben genannten Einsatzfeldern wird Eyetracking auch vielfach zur Erfassung verteilter Aufmerksamkeit bei der Mensch-Maschine-Interaktion eingesetzt. Im FVK-Bereich wurde Eyetracking beispielweise bereits in der Systementwicklung als Methode zur Validierung unterschiedlicher Arbeitssysteme verwendet. Konkret konnten die aufgezeichneten Blickdaten zur Analyse der Aufmerksamkeitsverteilung bei Lotsen angewendet werden (Herr & Herber, 2008). Es wurde beispielsweise die Informationsaufnahme der Lotsen in einem MSP-Team (Multi Sector Planning) im Vergleich zu einem ECPC-System (ein Executive Controller und ein Planer Controller) anhand der Analyse des Blickverhaltens eingehender untersucht. Die Ergebnisse geben Aufschluss über die Veränderungen für den Lotsen, die mit einer Umstellung auf MSP einhergehen würden.

² Weber (2002, S.82) definiert visuelle Beanspruchung bezüglich des Fluglotsenarbeitsplatzes als „die Auswirkung der durch die Augen aufgenommenen Informationen im Individuum, die zum Aufbau und zur stetigen Aktualisierung des Pictures erforderlich sind, das sich aus dem mentalen Gerätemodell (Flugsicherungssystem) und dem mentalen Geländemodell (dreidimensionaler Luftraum mit Flugzeugen und Flugbewegungen) zusammensetzt. Sie hängt auch von den individuellen Voraussetzungen und dem Zustand des Individuums ab“. Flugsimulationsstudien von Veltman und Gaillard (1996) haben gezeigt, dass der Lidschlag deutlich sensibler auf Belastungen des visuellen Systems bei der Informationsaufnahme als auf andere Formen mentaler Belastung reagiert.

1.4 Auswahlkriterien für Eyetracker und Fragestellung für die Untersuchungen

In den vorhergehenden Abschnitten wurden mögliche Einsatzgebiete aufgezeigt und kritische Punkte, wie die Notwendigkeit für eine möglichst nicht-invasive Implementierung, sowie der Bedarf für eine ausreichend hohe Messgüte aufgezeigt. Letzteres wurde hier eingehender untersucht, da es sich um eine Grundvoraussetzung handelt, die es zu erfüllen gilt, bevor über eine mögliche Anwendung der Technologie weitergehend nachgedacht werden kann.

Auf dem Markt gibt es aktuell ein großes Angebot unterschiedlicher video-basierter Eyetrackingsysteme (sog. videookulographische Systeme), die hier im Weiteren ausschließlich betrachtet werden. Für die Auswahl eines für die Untersuchung geeigneten Systems sind verschiedene Merkmale, wie z.B. die Applizierung der Sensorik in Betracht zu ziehen. Bisherige kopfbasierte Systeme, bei denen die Messsensorik an einem Kopfband oder einem Helm fixiert ist, haben den entscheidenden Nachteil, dass ihr Tragekomfort oftmals stark eingeschränkt ist. Nicht selten berichten Probanden nach kurzzeitigem Tragen von Kopfschmerzen aufgrund des zusätzlichen, ggf. sogar inhomogen verteilten Gewichtes auf dem Kopf (Lorenz et. al., 2004; Lorenz et al., 2007). Mittlerweile sind kontaktfreie Messgeräte auf dem Markt verfügbar, durch welche dieses Problem gelöst werden konnte. Für deren Einsatz im FVK- Bereich sollten jedoch einige weitere Kriterien erfüllt sein:

- Eine adäquate Blickerfassungsqualität sollte gegeben sein:
 - Die örtliche Auflösung sollte maximal sein, d.h. die Unterscheidbarkeit verschiedener Sehwinkel sollte so groß wie möglich sein.
 - Die zeitliche Auflösung, die die Frequenz bzw. die Anzahl der aufgenommenen Blickbewegungen und die Latenzzeit umfasst, sollte ausreichend hoch sein.
 - Die Fehlerrate bei der Datenerfassung sollte minimal sein.
- Das gewählte Verfahren muss robust gegenüber gelegentlichen Blickabwendungen des Benutzers sein.

- Schließlich ist darauf zu achten, dass die Blickregistrierung auch beim Tragen von Sehhilfen sicher und fehlerfrei funktioniert.

Insgesamt wurden drei Untersuchungen durchgeführt, die sich vor allem auf die Blickerfassungsqualität der drei unterschiedlichen Eyetrackingsysteme konzentrierten. Hierbei wurden die örtliche Auflösung bzw. die Genauigkeit des Seh winkels (hier im Folgenden auch als Winkelfehler bezeichnet) und die Fehlerrate, d.h. Datenverlustrate betrachtet. Die örtliche Genauigkeit ist von den jeweiligen Herstellern für jedes System angegeben. Für die vorliegenden Untersuchungen wurde der Winkelfehler (y) als Tangens des Quotienten aus Positionsabweichung (x) und dem mittleren Sehabstand (d), wie in der folgenden Formel dargestellt, berechnet:

$$y = \tan(x / d) \cdot 180 / \pi$$

Mittels der folgenden Untersuchungen wurde versucht, die von den Herstellern angegebenen Werte unter FVK-Bedingungen, d.h. am Lotsenarbeitsplatz zu replizieren.

2. Methode

Für alle Untersuchungen wurde ein quadratischer Standardarbeitsmonitor (503*503mm, 2048*2048 Pixel) aus dem FVK-Bereich verwendet. Die Sti-

51x/y	267x	483x	699x	915x	1131x	1347x	1563x	1779x	1995x
267y									
483y									
669y									
915y									
1131y									
1563y									
1779y									
1995y									

Abb. 2: Schematische Darstellung der verwendeten 10*10 Rasteraufteilung auf dem 2048*2048 Pixel Bildschirm.

muli wurden in einem Raster von 10*10 auf dem Bildschirm dargeboten (Abb. 2).

In den ersten beiden Untersuchungen wurden drei Stimuli (Quadrat, Kreuz und Kreis) vergleichbarer Größe (49 mm) verwendet. Um in der dritten Untersuchung zusätzlich noch den Effekt der Stimulusgröße zu kontrollieren, wurde gegen Ende dieser Untersuchung ein weiteres, kleineres Kreuz (25 mm) in einem zusätzlichen Durchgang dargeboten. In allen drei Untersuchungen wurden die Probanden zu Beginn über das experimentelle Setup informiert und instruiert, die dargebotenen Stimuli möglichst mittig zu fixieren,

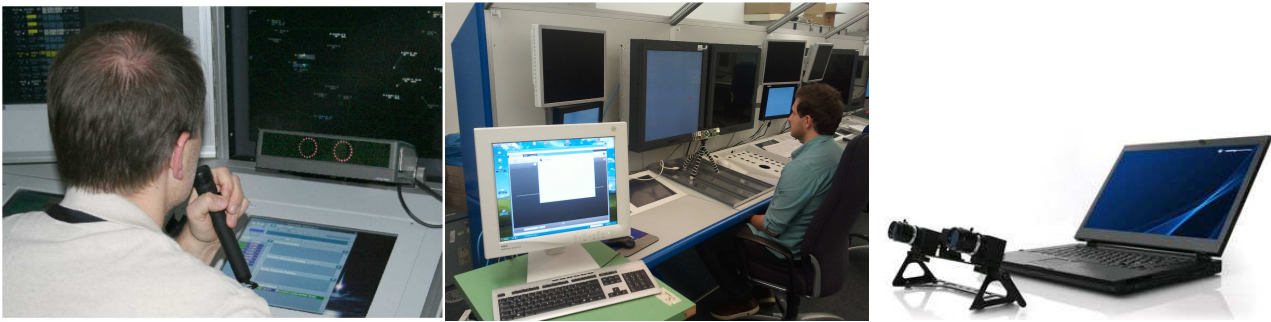


Abb. 3: Blickbewegungsmessung mit Tobii X120 (links), Tobii OEM (Mitte) und Face LAB5.0 (rechts, Quelle: Seeing Machines, 2011).

während diese im Sekundentakt (1 Sekunde Darbietungsdauer pro Position) von der linken oberen Ecke zur rechten unteren Ecke des Bildschirms und wieder zurück wanderten. Diese Prozedur wurde für jeden Stimulustyp durchgeführt.

In der ersten Untersuchung kam das Tobii X120 Eyetrackingsystem zum Einsatz (Abb. 3, links). Die vom Hersteller angegebene Messungsgenauigkeit beträgt 0.5° Sehwinkel. Die Untersuchung fand im Frühjahr 2010 an einer Controller Working Position (CWP) in einem der Radar-Simulatoren im Forschungszentrum der DFS statt. Nach Instruktion der Probanden erfolgten eine kurze Kalibrierung des Systems und schließlich die Durchführung des Versuchs, der ca. 10 Minuten dauerte.

Unter Verwendung einer OEM-Version des Tobii X120 Eyetrackingsystems (Abb. 3, Mitte), für das der Hersteller eine verbesserte Messgenauigkeit und geringere Ausfallrate angab, fand eine zweite Untersuchung im Herbst 2010, unter gegenüber der ersten Untersuchung vergleichbaren Bedingungen, statt. An dieser Untersuchung nahmen 22 Probanden teil.

In der dritten Untersuchung wurde das Eyetrackingsystem Face LAB 5.0 der Firma Seeing Machines verwendet (Abb. 3, rechts). Als Facial Feature Tracker bietet dieses System die Möglichkeit, sowohl Augenbewegungen als auch den Gesichtsausdruck zu erfassen. Die angegebenen Messungsgenauigkeit liegt laut Hersteller bei 0.5° - 1° Sehwinkel. An dieser Studie haben 27 Probanden, davon 24 Frauen, teilgenommen. 26 Probanden waren Psychologiestudenten der TU Darmstadt, eine Person war berufstätig. Der Versuch fand im November 2010 am Institut für Psychologie der TU Darmstadt statt. Eine Testeinheit

pro Versuchsperson dauerte ca. 15 bis 20 Minuten.

3. Ergebnisse

Abbildung 4 zeigt die Verteilung der örtlichen Genauigkeit auf den vorab definierten Positionen des Bildschirms in den drei Untersuchungen.

Die mittlere örtliche Genauigkeit des Tobii X120 lag bei 6.67° (Standardabweichung (SD) = 0.63). Das Eyetrackingsystem Tobii OEM erzielte eine mittlere örtliche Genauigkeit von 4.57° (SD = 0.32) und Face LAB 5.0 von 2.37° (SD = 0.74). Anhand des Verlaufs der Linien ist zu erkennen, dass das Tobii X120 die schlechteste Genauigkeit aufwies. Dagegen lieferte Face LAB5.0 die beste Genauigkeit. Der Mittelwert der Genauigkeit des Tobii OEM-Systems lag zwischen der der beiden anderen Systeme.

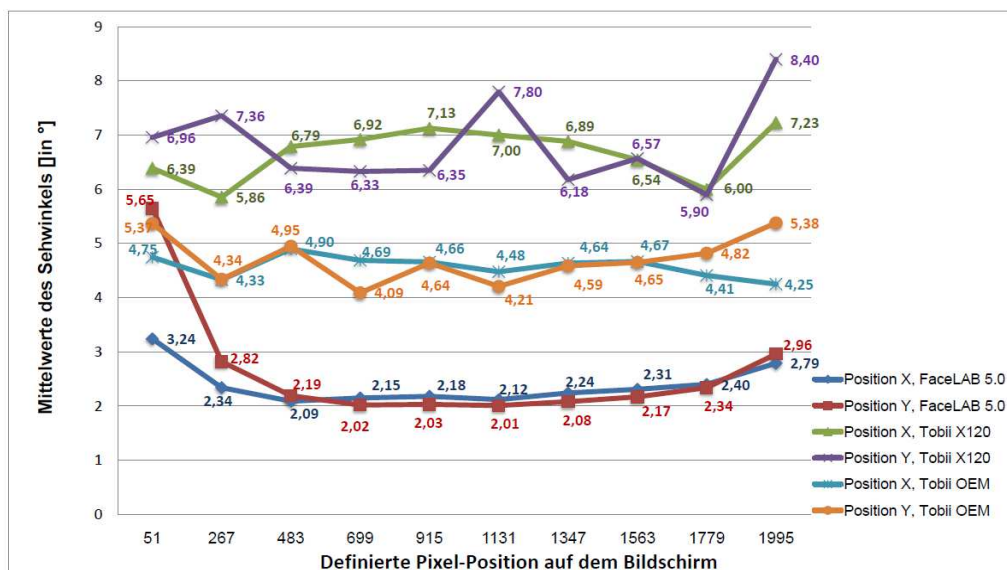


Abb.4: Mittelwert des Sehwinkels der drei Messgeräte (Face LAB5.0, TobiiX120 und Tobii OEM) auf die X und Y Bildschirmpositionen.

Eine detaillierte Aufschlüsselung der örtlichen Genauigkeit der drei Eyetrackingsysteme und für jeden Punkt auf dem dargebotenen 10x10 Stimulusraster ist Tabelle 1 zu entnehmen. Die grün markierten Werte waren kleiner oder gleich dem Mittelwert des Sehwinkels. Die gelb markierten Werte waren größer als der Mittelwert, aber kleiner als der doppelte Mittelwert. Die rot markierten Werte waren größer als der doppelte Mittelwert. Deutlich erkennbar ist, dass die Messgenauigkeit insbesondere an den Bildschirmrändern abfällt. Betrachtet man die Verteilung der Werte beim Face LAB5.0 System (s. Tab.1 un-

ten) so wird eine Art zentraler „Softspot“ deutlich. Allgemein scheint die Monitorgröße einen entscheidenden Einfluss auf die Messgüte zu haben.

Abbildung 5 zeigt die Verteilung des Datenverlustes auf die definierten Positionen des Bildschirms der drei Systeme. Im Mittelwert hatte das Tobii X 120 System einen Datenverlust von 32% ($SD = 0,08$), Tobii OEM von 4% ($SD = 0,01$) und Face LAB5.0 von 7% ($SD = 0,09$) aufzuweisen. Auch hier scheint sich der Eindruck einer deutlich schlechteren Messung an den Bildschirmrändern zu bestätigen.

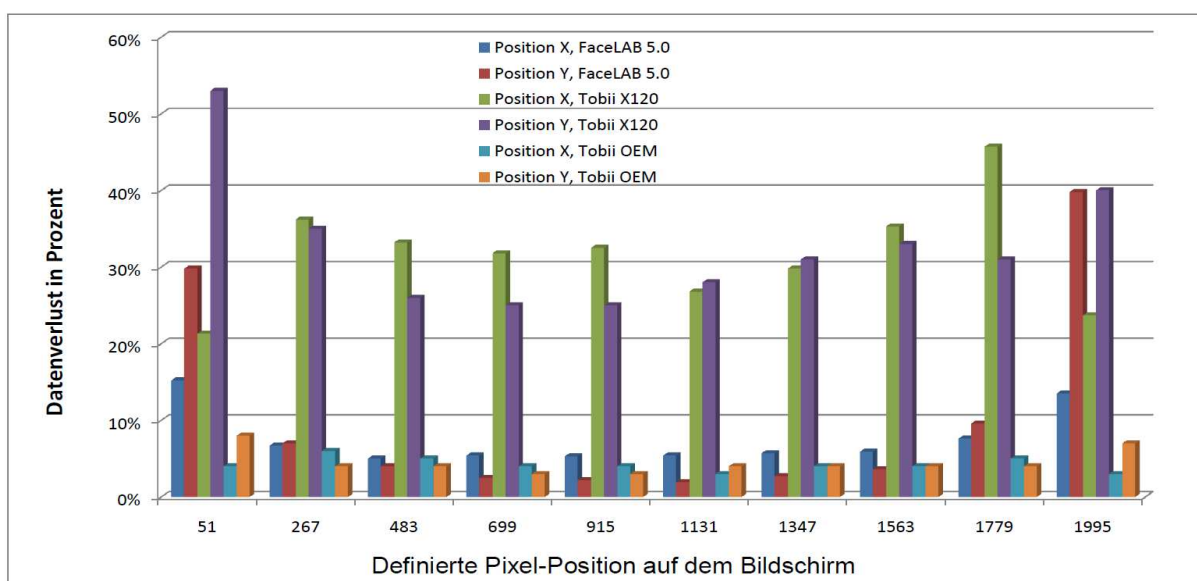


Abb. 5: Verteilung des Datenverlusts der drei Messgeräte (Face LAB 5.0, Tobii X120 und Tobii OEM) auf die verschiedenen Positionen des Bildschirms

Pixel-Position Y	Pixel-Position X (Tobii X120, $\bar{x} = 6,67^\circ$)									
	51	251	451	651	851	1051	1251	1451	1651	1851
51	6,68	6,41	6,88	6,94	7,05	6,98	6,92	6,75	6,48	7,1
251	6,87	6,61	7,07	7,14	7,24	7,18	7,12	6,95	6,68	7,29
451	6,59	6,12	6,59	6,65	6,76	6,69	6,64	6,47	6,19	6,81
651	6,36	6,09	6,56	6,63	6,73	6,67	6,61	6,44	6,16	6,78
851	6,76	6,1	6,57	6,64	6,74	6,68	6,62	6,45	6,17	6,79
1051	7,1	6,83	7,3	7,36	7,47	7,4	7,35	7,17	6,9	7,52
1251	6,64	6,02	6,48	6,55	6,65	6,59	6,53	6,36	6,09	6,71
1451	6,48	6,21	6,68	6,74	6,85	6,79	6,73	6,56	6,28	6,9
1651	6,19	5,88	6,34	6,41	6,51	6,45	6,39	6,22	5,95	6,56
1851	7,2	7,13	7,59	7,66	7,76	7,7	7,64	7,47	7,2	7,81
Pixel-Position Y	Pixel-Position X (Tobii OEM, $\bar{x} = 4,57^\circ$)									
	51	251	451	651	851	1051	1251	1451	1651	1851
51	6,08	5,99	5,95	5,51	5,00	4,87	5,38	5,55	4,78	3,52
251	4,54	3,92	4,80	4,62	4,20	3,71	3,58	4,23	4,48	4,41
451	5,30	5,40	5,32	5,19	4,86	4,90	4,96	5,10	3,96	3,75
651	3,98	2,99	4,35	4,48	4,01	3,98	3,95	4,18	4,59	4,87
851	5,18	4,93	4,84	4,95	4,67	4,80	4,96	5,09	3,67	3,56
1051	3,83	3,11	4,64	4,96	4,70	4,30	4,17	4,32	4,49	4,60
1251	4,81	4,86	4,77	4,73	4,76	5,20	5,07	5,19	3,60	3,73
1451	4,21	3,69	5,10	5,10	4,68	4,92	4,60	4,68	4,83	4,95
1651	5,18	4,85	4,97	5,04	5,05	5,43	5,39	5,36	3,83	4,05
1851	4,55	4,38	5,63	5,58	5,23	5,01	5,04	5,67	5,38	5,92
Pixel-Position Y	Pixel-Position X (Face LAB5.0, $\bar{x} = 2,37^\circ$)									
	51	267	483	699	915	1131	1347	1563	1779	1995
51	8,41	4,33	3,16	3,68	5,17	4,99	6,33	6,45	7,12	7,94
267	4,47	2,94	2,28	2,63	2,68	2,62	2,58	2,52	2,73	2,93
483	3,13	2,21	2,16	2,15	1,93	2,03	1,91	2,00	2,04	2,53
699	2,89	2,08	2,02	1,85	1,84	1,59	1,80	1,90	1,99	2,36
915	3,01	2,22	1,90	1,82	1,77	1,73	1,89	1,94	1,87	2,19
1131	2,51	2,08	1,97	1,88	1,90	1,88	1,78	1,90	1,98	2,21
1347	2,54	2,04	1,86	1,92	1,92	1,90	1,98	2,03	2,17	2,46
1563	2,58	2,09	1,87	2,06	1,92	2,02	2,05	2,23	2,20	2,78
1779	2,87	2,20	1,97	1,99	2,18	2,04	2,28	2,52	2,51	2,97
1995	2,99	2,84	2,83	3,04	2,72	2,32	2,70	2,95	3,38	3,94

Tab. 1: Überblick der Verteilung des Winkelfehlers (in Grad Sehwinkel) auf den definierten Pixel-Positionen auf dem Bildschirm von Tobii X120, Tobii OEM und Face LAB5.0. Grüne Markierung: Werte $\leq$ Mittelwert des Seh winkels, gelbe Markierung: Mittelwert $\leq$ Werte $<$ doppelter Mittelwert, rote Markierung: Werte $>$ doppelter Mittelwert.

4 Diskussion und Ausblick

4.1 Diskussion

Aufbauend auf den hier vorgestellten Ergebnissen ist zu erkennen, dass Face LAB 5.0 im Vergleich zu den anderen beiden Systemen noch die beste örtliche Genauigkeit (Abweichung: 2,37°, Datenverlustrate: 4%) lieferte. Das Tobii OEM (s. Tabelle 2) System hatte hingegen die niedrigste Datenverlustrate. Das Tobii X120 System zeigte im relativen Vergleich sowohl bei der örtlichen Genauigkeit als auch beim Datenverlust die schlechtesten Ergebnisse. Insgesamt betrachtet lassen die vorliegenden Ergebnisse nur eingeschränkte Anwendungsmöglichkeiten für Eyetrackingsysteme im FVK-Bereich erkennen. Die erzielte Genauigkeit und Datenverlustrate sind keinesfalls hinreichend, um weitere Schritte in Richtung einer prototypischen Anwendung als unterstützendes Eingabemedium anzustreben.

Insbesondere an den Randpositionen des großformatigen, quadratischen ATC-Standardbildschirms erwies sich die Blickerfassung in Bezug auf die hohen Fehlerraten als besonders problematisch. Ein solcher Monitor erzwingt sehr große vertikale Augenbewegungen, um den oberen oder unteren Bildschirmrand zu betrachten. Insbesondere bei Vertikalbewegungen nach oben verschwinden Teile der Iris hinter dem oberen Augenlid. Die auf die Erfassung des runden Iris-musters ausgelegten Blickbewegungsmessgeräte können unter diesen Bedingungen keine vollflächige Iris detektieren und folglich keine Messwerte liefern.

Wie sich in den vorliegenden Untersuchungen gezeigt hat, besteht eine weitere Schwierigkeit in der erforderlichen Expertise bei der Handhabung eines Eyetrackingsystems. Der Einsatz bedarf ausreichend geschulten Personals. Ein erhöhter Aufwand entsteht auch durch die für jeden Probanden erforderliche Kalibrierung, die vor jedem Arbeitsgang neu ausgeführt werden muss. Ein solcher Aufwand ist schwer zu argumentieren und spricht ebenfalls eher gegen den unmittelbaren Einsatz dieser Technologie zum aktuellen Stand.

Zusammenfassend hat das kontaktlose Eyetracking große Vorteile aufgrund seiner Nicht-Invasivität. Akzeptanzprobleme bei den Probanden werden dadurch im Gegensatz zu kopf-/helmbasierten Systemen von vorne herein minimiert. Zudem kann das kontaktlose System relativ problemlos im Hintergrund mitlaufen und wird im Idealfall dann irgendwann im Verlauf der Untersuchung gar nicht mehr bemerkt. Durch diese Technologie ist eine kontinuierliche Datenaufzeichnung möglich, ohne dass eine andere Tätigkeit für die Messung unterbrochen werden muss, wie dies z.B. bei einem Fragebogen der Fall ist.

Andererseits erscheint aufgrund der momentan fehlenden technischen Voraussetzungen der Einsatz von Blicksteuerung zur Kontrolle von Benutzungsschnittstellen an FVK-Arbeitsplätzen in den nächsten Jahren nur schwer umsetzbar. Für einen sinnvollen Einsatz müsste die Blickerfassungsqualität des Eyetrackings grundlegend verbessert werden. Vielversprechende Ansätze

Zu erreichende technische Eigenschaften		Tobii		Face LAB5.0
		X120	OEM	
Örtliche Auflösung	Mittelwert örtlicher Genauigkeit	6,67°	4,57°	2.37°
	Mittelwert der Fehlerrate	32%	4%	7%
Zeitliche Auflösung	Frequenz (Frames/ sec.)	60	30	60
Blickregistrierung mit Sehhilfe		möglich		möglich
Blickerfassung am Bildschirmrand		schlecht	zufriedenstellend	befriedigend

Tab. 2: Überblick des Vergleichs zwischen Tobii- und Face LAB5.0 Eyetrackingsystemen

liegen in dem Versuch der Kopplung von Systemen, um eine höhere Abdeckung des Bildschirms und insbesondere des Bildschirmrands zu erreichen. Notwendig ist jedoch auch dann, dass die Bildauswertelgorithmen robust gegenüber eintretenden Verdeckungen der Iris bei starken vertikalen Blickbewegungen ausgelegt werden. Hier ist jedoch noch substantielle Grundlagenforschung und Entwicklungsarbeit zu betreiben. Geklärt werden müsste unter anderem z.B. auch wie hoch die Robustheit der gewählten Verfahren gegenüber gelegentlichen Blickabweichungen des Benutzers ist.

4.2 Ausblick

Mit der in den vorliegenden drei Untersuchungen festgestellten Messgüte des Eyetrackers ließen sich voraussichtlich folgende Anwendungsmöglichkeiten realisieren:

- **Kapazitätsanalyse**

Zur Erfassung der Kapazität eines ATM-Systems wurden zwei standardisierte Methoden von Eurocontrol entwickelt: die Capacity Analysis Methode (CAPAN) und die Integrated Task Analysis Methode (ITA). Beide Methoden haben ihre Schwächen: CAPAN ist nur geeignet für bekannte, seit längerem eingesetzte ATM-Systeme, da das aufgabenorientierte Zeitschema für die Arbeit mit einem zukünftigen System a priori nicht bekannt ist (Bierwagen & Tautz, 2006). ITA basiert auf Beobachtungen und Einschätzungen von Experten. Da sich kognitive Verarbeitungsprozesse nur schwer durch Beobachtung erfassen lassen, wird diese Methode häufig kritisiert (Bierwagen & Tautz, 2006). Eyetracking könnte in diesem Kontext als Unterstützung eingesetzt werden, da die Blickbewegungsanalyse den Vorteil hat, unbewusste Vorgänge der Aufmerksamkeitsverteilung zu registrieren.

- **Analyse von Blickverhaltensmustern**

Die Art der Tätigkeit von Fluglotsen wird überwiegend als Informationsverarbeitung bzw. geistige Arbeit definiert, welche die verschiedenen Informationsverarbeitungsschritte: Informationsaufnahme (Wahrnehmen und Entdecken), Informationsverarbeitung (Erkennen und Entscheiden) und Informationsabgabe (Handeln) beinhaltet (s. hierzu auch Eurocontrol, 1996). Bei der geistigen Arbeit ist immer

schwer zu beobachten, wann, wie und welche Verarbeitungsprozesse stattfinden. In bisherigen Blickbewegungsanalysen kommen Forscher zu der Schlussfolgerung, dass sich jede Teilhandlung der geistigen Arbeit teilweise unterschiedlichen Ausprägungen der Blickbewegung zuordnen lässt. So spiegelt die Anzahl der Fixationen die Suche nach der Lösung von Aufgaben wider. Die relative Verweildauer auf einem sog. „Area of Interest“ kann als Maß für die Bedeutsamkeit eines Objekts angesehen werden. Ferner gibt es Hinweise darauf, dass über die Richtung der Augenbewegungen (nach links oder nach rechts) ein Bezug zur bevorzugten Problemlösetechnik einer Person hergestellt werden kann, da weithin eine Bewegung bei Kopfrechenaufgaben nach links auf eine Visualisierung und eine Bewegung nach rechts auf eine abstrakte Lösung des Problems hindeutet (Oster & Stern, 1980).

Die Fragen, die sich bisher nur teilweise mittels Aussagen der Lotsen oder durch Beobachtung von Experten beantworten ließen, könnten mit Hilfe der Blickbewegungsanalyse objektiv untersucht werden: z.B. wie viel Zeit braucht ein Lotse, um einen Konflikt zu erkennen und zu lösen; welche Informationen auf dem Hauptbildschirm und ggfs. welche Informationen werden auf weiteren Bildschirmen benötigt, um eine angemessene Entscheidung treffen zu können? Welche Problemlösestrategie hat ein Lotse angewendet? Es erscheint als eine interessante Fragestellung, nicht zuletzt auch aufgrund des Potenzials für die Ausbildung zukünftiger Lotsen, zu prüfen, inwieweit die Blickpfadanalyse geeignet wäre, um unterschiedliche kognitive Stile zu identifizieren. Effiziente Lösungsstrategien könnten auf diese Weise in Trainingsprogramme und Schulungen einfließen.

- **Systemgestaltung**

Aufgrund des direkten Zusammenhangs zwischen Informationsdarstellung am Bildschirm und der Belastung des visuellen Informationsverarbeitungskanals lassen sich mittels Blickbewegungsanalyse die Art der Informationsdarstellung und das Displaydesign einer Anwendung objektiv bewerten und entsprechend optimieren.

■ Erfassung des mentalen Zustandes

In diesem Zusammenhang wurden bereits die mentale und visuelle Beanspruchung sowie Vigilanz genannt, allerdings lässt sich die mentale Beanspruchung nicht sonderlich präzise durch die Blickbewegungen abbilden. Andererseits scheint der Lidschlag ein geeigneter Indikator für die Vigilanzanalyse und die visuelle Beanspruchung zu sein. Ein denkbarer Einsatz wäre Untersuchungen zum Vigilanzverlust, um die Gegenmaßnahme besonders im Nachtbetrieb geben zu können.

Weitere vielversprechende Einsatzfelder, wie z.B. als Aufmerksamkeitsunterstützung für ein Konflikterkennungssystem sind denkbar, sobald eine ausreichende Messgüte der Systeme gegeben ist.

Jedoch bleiben nach wie vor eine Reihe offener Fragen. Da Flugverkehrskontrolle von Teamarbeit geprägt ist, bleibt es zu klären, ob Eyetracking-Messgeräte generell in der Lage sind, mehrere Benutzer gleichzeitig vor einem Gerät zu identifizieren und voneinander zu unterscheiden. Kommerzielle Systeme aus dem Bereich der Unterhaltungsindustrie wie z.B. Microsofts Kinect scheinen derartige Hürden bereits gemeistert zu haben. Inwieweit bestehende Lösungen genutzt werden können, um die o.g. offenen Fragen adressieren zu können, wird die weitere Forschung zeigen müssen. Auch wird sich zeigen, ob für diese Technologie nach jahrzehntelangem Einsatz in der Grundlagenforschung ausreichende Voraussetzungen gegeben sind, um den Sprung in die Anwendung zu schaffen.

Literaturverzeichnis

Bach-Y-Rita, P., & Kercel, S. W. (2003). Sensory substitution and the human-machine interface. *Trends in Cognitive Science*, 541-546.

Beatty, J. (1982). Task-evoked papillary response, processing load and the structure of processing. *Psychological Bulletin*, 91, 276-292.

Bierwagen, T., Tautz, A. (2006). Methodik zur experimentellen Bestimmung des Kapazitätssteigerungspotenzials neuer Air Traffic Management Systeme. *TE im Fokus* (1/10), 2-10. Langen: Deutsche Flugsicherung.

Birren, J.E., Casperson, R.C. & Botwinik, J. (1950). Age Changes in Pupil Size. *Journal of Gerontology*, 5, 216-221.

Bouma, H. (1962). Size of the Static Pupil as a Function of Wave-Length and Luminosity of the Human Eye. *Nature*, 193.

Bösel, R. (1987): *Physiologische Psychologie*. Berlin, New York, De Gruyter.

Drewes, H. (2010). *Eye Gaze Tracking for Human Computer Interaction*. Dissertation, Ludwig-Maximilians-Universität München.

Esser, P. & Bos, T.J.J.(2006a). Eye Gaze interactive ATC workstation. *Report NLR-TP-2006-735. National Aerospace Laboratory NLR*. Amsterdam: Universität Maastricht.

Esser, P. & Bos, T.J.J.(2006b). Development and Evaluation of an Eye Gaze interactive Air Traffic Control workstation. Amsterdam: Universität Maastricht. Zugriff am 15.06.201: <http://inoworkshop.eurocontrol.fr/Previous/index-29116.pdf>

Eurocontrol (1996). Proceedings of the First EUROCONTROL Human Factors Workshop Cognitive Aspects in ATC. S.14. Brüssel.

Galley, N. (2001). Physiologische Grundlagen, Meßmethoden und Indikatorfunktionen der okulomotorischen Aktivität. In Rösler, F.(Hrsg.), *Grundlagen und Methoden der Psychophysiologie* (S. 237-316). Göttingen: Hogrefe.

Gambill, H.D., Dogle, K.A. & Kearns, T.P. (1967). Mydriatic Effect of four Drugs Determined with Pupillograph. *Arch Ophthal*, 77, 740-746.

Grandt, M., Hagemann, K., Herr, S., & Teichmann, M. (2009). Blickbewegungsmessung und Blicksteuerung in der Flugsicherung - Quo Vadis?. *TE im Fokus* (1/10), 2-10. Langen: Deutsche Flugsicherung.

Herr, S., & Herber, A. (2008). Zusammenfassung der Ergebnisse im Projekt MSP-D/L. *TE im Fokus* (2/08), 17-23. Langen: Deutsche Flugsicherung.

Hornung, J. (1967). Pupillenbewegung nach einem Sprung der Reizlichtintensität. *Pflügers Archiv*, 296, 39-48.

Hyönä, J., Tammola, J., & Alaja, A.-M. (1995). Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. *The Quarterly Journal of Experimental Psychology*, 48A(3), 598-612.

Jacob, R. (1990). What you look at is what you get. *CHI '90 Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people* (S. 11-18). New York, NY, USA: ACM.

Jacob, R., & Karn, K. (2003). Eyetracking in Human-Computer Interaction and usability research: Ready to deliver the promises. In R. H. J. Hyönä, *The mind's eye: Cognitive and applied aspects of eye movement research* (S. 573-605). Amsterdam: Elsevier.

Just, M., & Carpenter, P. (1993). The intensity dimension of thought: Pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology*, 47(2), 310-339.

Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, USA: Prentice-Hall.

Krüger, F. (2000). *Coding of temporal relations in semantic memory: Cognitive load and task-evoked pupillary response*. Münster: Waxman.

Kumnik, L.S. (1954). Pupillary Psychosensory Restitution and Aging. *Journal of the Optical Society of America*, 44, 735-741.

Lorenz, B., Biella, M., Schmerwitz, S., Többen, H. (2004). Verlässlichkeit der Interaktion zwischen Pilot und Assistenzsystem: Erfahrung mit seltenen kritischen Ereignissen in Simulationsszenarien. In M. Grandt (Hrsg.), *Verlässlichkeit der Mensch-Maschine-Interaktion (DGLR-Bericht 2004-03)*, S.271-294. Bonn: Deutsche Gesellschaft für Luft- und Raumfahrt.

Lorenz, B., Biella, M., Teegen, U., Stelling, D., Wenzel, J. et al. (2007). Performance, situation awareness and visual scanning of pilot receiving onboard taxi navigation supporting during simulated airport surface operation. *Human Factors and Aerospace Safety*, 6, S.135-154.

Oster, P.J. & Stern, J.A. (1980). Measurement of eye movements-Electrooculography. In I. Martin & P.H. Venables (Hrsg.), *Techniques in psychophysiology*. Chichester: Wiley.

Schleicher, R., Galley, N., Briest, S., & Galley, L. (2008). Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired? *Ergonomics*, 71, 982-1010.

Schultheis, H. (2004). *Pupillengröße und kognitive Belastung*. Diplomarbeit. Saarbrücken: Universität des Saarlandes, Institut für Psychologie.

Seeing Machines (2011). Zugriff am 05.06.2011: <http://www.seeingmachines.com/>

Siebert L. E., & Jacob R. J. K. (2000). Evaluation of Eye Gaze Interaction. In *Proceedings of CHI 2000: Human Factors in Computing Systems*. The Hague, The Netherlands: ACM Press.

Steinhauer, S.R., Siegle, G.J., Condray, R. & Pless, M. (2004). Sympathetic and parasympathetic innervations of papillary dilation during sustained processing. *International Journal of Psychophysiology*, 52, 77-86.

Veltman, J. A. & Gaillard, A. W. (1996). Physiological indices of workload in a simulated flight task. *Biological Psychology*, 42, 323-342.

Weber, P. (2002). *Die Lidschlagrate als Vergleichsparameter für aktuelle und zukünftige Flugsicherungssysteme hinsichtlich Ermüdung und visueller Beanspruchung von Piloten*. Berlin: Verlag dissertation.de.

Abkürzungsverzeichnis

ATC	Air Traffic Control
CWP	Controller Working Position
FKV	Flugverkehrskontrolle
OEM	Original Equipment Manufacturer
MMI	Mensch-Maschine-Interaktion
SD	Standabweichung

Kopplung zwischen einem Schnellzeitsimulationstool und betrieblichen Planungssystemen

Planung und Durchführung einer automatisierten Simulation im Rahmen von iPort

Katharina Schwentek, TE/AT

1. Einleitung

Die optimale Ausnutzung verfügbarer Kapazitäten an Hub-Flughäfen spielt mit dem wachsenden Luftverkehrsaufkommen eine immer bedeutendere Rolle. Die Weiterentwicklung bereits eingesetzter Planungssysteme sowie Untersuchung neuer Ansätze zur Unterstützung der Verkehrsabwicklung wurde in den letzten Jahren intensiviert, um neue Wege zu finden, die Leistungsfähigkeit des Luftverkehrs zu steigern. Je komplexer sich die Situation gestaltet, desto bedeutender sind technische Assistenzsysteme, um den Lotsen zu entlasten und ihn in seinen Aufgaben proaktiv zu unterstützen.

In Europa vergibt die CFMU (Central Flow Management Unit) Abflugslots für eine kapazitive Vorsteuerung des Luftverkehrs, wenn in bestimmten Sektoren oder an den Zielflughäfen Engpasssituationen zu erwarten sind. Dabei wird in drei Phasen unterschieden: strategisch (sechs Monate bis zwei Tage vor Flugereignis), prätaktisch (zwei Tage vor Flugereignis) und taktisch (Tag des Flugereignisses). Neben dieser europaweiten Betrachtung des Verkehrs werden lokal an Hub-Flughäfen verschiedene taktische Planungssysteme eingesetzt, die die Nachfragesituation der kommenden Stunde betrachten. Sogenannte Arrival Manager (AMAN) erstellen eine Sequenz für den Anflugverkehr auf eine Landebahn, Departure Manager (DMAN) führen die Sequenzplanung auf der Abflugseite durch. Seit einigen Jahren ist an großen Flughäfen der sogenannte A-CDM (Airport Collaborative Decision Making) Prozess eingeführt, der den Datenaustausch und die Kommunikation zwischen allen beteiligten Partnern (Flughafen, Flugsicherung, Airlines) verbessern soll.

Neben bereits betrieblich genutzten Systemen werden auch Prototypen neuer Systeme entwickelt und deren Verwertbarkeit für die Praxis untersucht. Ein Beispiel hierfür ist der CLOU/FMAN (Cooperative Local Resource Planner / Flow Manager), ein Planungssystem, das lokal an einem Flughafen für die kommenden Stunden die De-

mand- und Kapazitätsprognose betrachtet und mögliche Engpässe sowie Optimierungsvorschläge aufzeigt. Es wird als sogenanntes prätaktisches Planungssystem bezeichnet, ist jedoch in seiner Funktionsweise von der CFMU zu unterscheiden. Während die CFMU den gesamten europäischen Luftverkehr betrachtet, liegt der Fokus von CLOU/FMAN lokal auf einem Flughafen.

Im Rahmen der Projekte¹ K-ATM (Kooperatives ATM), WFF (Wettbewerbsfähiger Flughafen) sowie iPort (Innovativer Airport) lag der Fokus zunächst auf der Weiterentwicklung von Einzelsystemen, die den taktischen Planungshorizont (AMAN, DMAN) bzw. prätaktischen Planungshorizont (CLOU/FMAN) betrachten ([2], [3], [4]). In Zusammenarbeit mit dem operativen Betrieb wurden neue Funktionalitäten implementiert und getestet, um die Systemverwertbarkeit zu steigern.

Im Rahmen von iPort sollen die taktischen und prätaktischen Planungen sinnvoll miteinander vereint werden, um eine übergreifende Steuerung des Verkehrs an einem Hub-Flughafen zu ermöglichen. Mit der Kopplung von prätaktischen und taktischen Planungssystemen soll untersucht werden, welche Effekte sich durch eine aufeinander abgestimmte prätaktische Mengenvorsteuerung des Arrival- und Departureverkehrs in Hinblick auf die Verkehrsabwicklung und Kapazitätsausnutzung erreichen lassen.

2. Beteiligte Planungssysteme

2.1 Prätaktische Planungssysteme

CLOU/FMAN als prätaktisches System analysiert die Nachfrage- und Kapazitätsprognose für die kommenden zwei bis sieben Stunden an einem Hub-Flughafen. Bei potentiellen Engpässen optimiert der Optimierungskern, der sogenannte

¹ durch das Bundesministerium für Wirtschaft und Technologie geförderte Projekte im Luftfahrt-Forschungsprogramm

FMAN, die Bahnnutzungsstrategie (Betriebsverfahren) und bestimmt eine optimierte Verteilung der An- und Abflugmengen auf den einzelnen Start- und Landebahnen. Dabei werden betriebliche Vorgaben wie Minstdauer eines Betriebsverfahrens (BV) oder erlaubte An- und Abflugrouten je Bahn berücksichtigt. Somit wird in der Bestimmung der optimierten Menge an Ankünften auf einer Bahn gleichzeitig die Situation auf der Gegenseite, der Abflugseite berücksichtigt und analog andersherum.

Im Ergebnis kann zwischen unterschiedlichen Optimierungsstufen gewählt werden (vgl. Abbildung 1). In der ersten Stufe wird das heute im operativen Betrieb angewandte Prinzip First Come First Serve FCFS restriktiv bei der Mengenermittlung berücksichtigt. In den weiteren beiden Optimierungsstufen wird diese Restriktion aufgeweicht (FCFS nur in den einzelnen Arrival- bzw. Departure-Strömen) oder ganz außen vorgegeben (zur Ermittlung der maximalen Verbesserung der Verkehrsabwicklung).

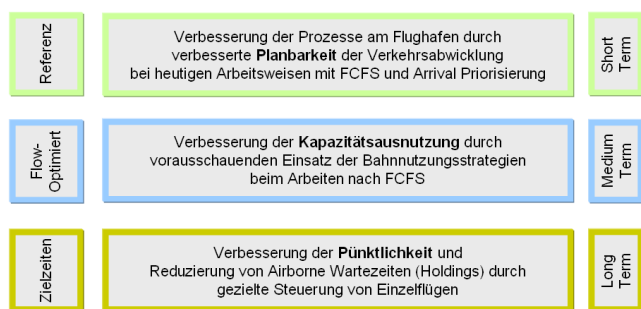


Abb. 1: Optimierungsstufen des FMAN

2.2 Taktische Systeme

AMAN erstellt für den Anflugverkehr auf eine Landebahn eine Arrival-Sequenz. Dabei wird mit dem sogenannten *AMAN-Flow* als Grundlage für die Separation gearbeitet. Der *AMAN-Flow* wird im operativen Betrieb vom Lotsen manuell eingestellt und entspricht dem durchschnittlichen Abstand in [NM] zwischen zwei aufeinander folgenden Ankünften, wobei der Abstand bei einzuhalten der Wirbelschleppenstaffelung ggf. größer ausfallen kann. Über den *AMAN-Flow* ermittelt das System für jeden Anflug eine Target Initial Approach Fix Time (TAFT), eine Zielzeit über dem Metering Fix (MF), von der die Target Landing Time (TLDT) abgeleitet wird.

DMAN erstellt für den Abflugverkehr auf einer Startbahn eine optimierte Departure-Sequenz. Dabei wird unter Berücksichtigung der Abflugfixe und dem damit verbundenen Mindestabstand zwischen zwei Abflügen auf derselben Standard Instrument Route (SID) für jeden Abflug eine sogenannte Target Startup Approval Time (TSAT) berechnet, d.h. eine Zielzeit zum Anlassen der Triebwerke. Über die Rollzeiten zur Schwelle, die als Standardwerte für jede Position hinterlegt sind, wird die Target Take-off Time (TTOT) ermittelt.

Im Rahmen der automatisierten Simulation wird der DFS-eigene generische *AMAN Frankfurt* als Prototyp (Version noch nicht operativ im Einsatz) sowie der Departure Manager *darts* von der Firma delair verwendet.

2.3 Definition der Planungshorizonte

In der automatisierten Simulation kommen der Arrival Manager AMAN und der Departure Manager darts als sogenannte taktische Systeme und der CLOU/FMAN als prätaktisches Planungssystem zum Einsatz. Die Systeme besitzen unterschiedliche Planungshorizonte, die sinnvoll miteinander verknüpft werden müssen:

AMAN: AMAN hat einen Planungshorizont von 100 NM um jeden MF. Alle Arrivals, die in diesen Horizont einfliegen, werden in der Sequenzplanung berücksichtigt und fallen für den AMAN somit in seinen taktischen Planungshorizont.

Darts: Die Planung eines Departures beginnt in darts ab [TOBT – 40min]. D.h. alle Departures, deren Estimated Off-block Time (EOBT, aus Flugplan) 40 min vor der aktuellen Zeit liegt, werden von darts in der Sequenzplanung berücksichtigt.

CLOU/FMAN: Der Planungshorizont von CLOU/FMAN kann variabel zwischen 2 und 7 Stunden eingestellt werden. Das System plant von der aktuellen Zeit an bis zum Ende des Planungshorizonts und berücksichtigt dabei alle Flüge, deren Schedule oder Estimate Times innerhalb dieses Horizonts liegen.

Die Definition, ab welcher Vorlaufzeit die taktischen Planungen priorisiert und von den CLOU/FMAN Planungen nicht mehr beeinflusst werden sollen, wird als variabler Parameter gestaltet, um die prätaktisch-taktische Kopplung von vornherein nicht zu statisch aufzubauen. Als

Standardwert wird angenommen, dass CLOU/FMAN die von den taktischen Systemen verplanten Flüge ab einer Vorlaufzeit von 30 min nicht mehr verändern darf (vgl. Abbildung 2). Die Variation kann in 10 min Schritten bis maximal zum eingestellten CLOU/FMAN Planungshorizont variiert werden.

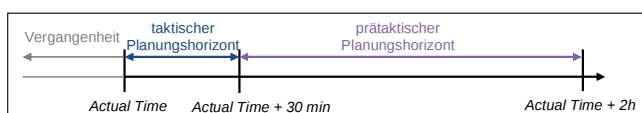


Abb. 2: Definition prätaktischer und taktischer Planungshorizont

3. Ziel der automatisierten Simulation

Gegenstand der Untersuchung des Teilpakets AP1.3.2 in iPort ist es, die Sequenzplanungen der taktischen Systeme AMAN und darts mit der Mengenplanung des prätaktischen Systems CLOU/FMAN zu koppeln. Dabei ist das übergeordnete Ziel, den potentiellen Nutzen zu analysieren, den die Mengenvorsteuerung für die Abwicklung des Verkehrs an einem Hub-Flughafen haben kann. Das Schnellzeit-Simulationstool AirTOP dient dabei als Werkzeug zur Umsetzung einer Simulation. Eine Realzeitsimulation mit Lotsen und Piloten ist in diesem Teilprojekt nicht vorgesehen.

Die Ergebnisse werden dafür ausschlaggebend sein, wie mit dem Systemverbund weiter verfahren wird. Diese Untersuchung soll zeigen, ob eine weitere Entwicklung der prätaktischen Mengenvorsteuerung und die Integration dieses neuen Ansatzes in die operative Systemlandschaft einen positiven Beitrag zur effizienten Verkehrsabwicklung liefern kann.

3.1 Etablierung eines Systemverbunds

Im Folgenden wird beschrieben, welche Aufgaben jedes der beteiligten Systeme im Verbund hat und welche Erwartung damit verknüpft wird.

AirTop als Schnellzeitsimulator ist für die Durchführung der Verkehrsabwicklung zuständig. Zunächst stellt das System den Datenlieferanten für alle gekoppelten Systeme dar. AirTop übermittelt die Flugpläne sowie Updates davon. AMAN benötigt zudem Trackdaten im CAT62-Format (Luft-raum) und darts Trackdaten im CAT11-Format (Bodenlage). Damit die Effekte der prätaktischen

Vorsteuerung auf die taktische Planung untersucht werden können, muss AirTOP in der Lage sein, die Zielzeiten von AMAN und DMAN bestmöglich umzusetzen. Nach Durchführung der Simulation können die für dieses Projekt relevanten Auswerteparameter (vgl. Kapitel 8) aus AirTOP herausgezogen werden.

CLOU/FMAN berechnet seine optimierte Lösung der Bahnnutzungsstrategie. Im Ergebnis kommt für jede Start- und Landebahn eine optimierte Menge an Arrivals sowie Departures je 10 min heraus. Dies steht für alle drei Optimierungsstufen (vgl. Kapitel 2.1) und den gewählten Planungshorizont (max. 7 Stunden) zur Verfügung.

Die optimierte Arrival-Menge wird je Runway und Zeitintervall an den **AMAN** gesendet, der daraus wiederum seinen *AMAN-Flow* berechnet. Um die Planungen stabil zu halten, werden zwei Werte übermittelt. Zum einen wird die maximal erlaubte Anzahl an Arrivals pro 10 min entsprechend des ausgewählten Betriebsverfahrens übermittelt. Dieser „*Flow Capacity*“-Wert ändert sich nur dann, wenn sich auch die Bahnnutzungsstrategie ändert. Der „*Flow Rate*“-Wert hingegen kann sich alle 10 min ändern. Er gibt die Anzahl an Arrivals an, die von CLOU/FMAN tatsächlich je Intervall verplant wurden. In Engpassssituationen auf Arrival-Seite wird *Flow Capacity* gleich der *Flow Rate* sein. Die *Flow Rate* wird von AMAN bestmöglich umgesetzt, indem er aus der Mengenvorgabe [ARR/10 min] seinen *AMAN-Flow* [NM] je Landebahn berechnet. Es ist also möglich, dass AMAN je 10 min unterschiedlich viele Arrivals verplant, der *AMAN-Flow* bleibt dabei jedoch konstant und ist ein Indiz dafür, ob gerade Arrivals oder Departures priorisiert werden sollen. Folgende Abbildung 3 veranschaulicht die Flow-Übermittlung CLOU-AMAN:

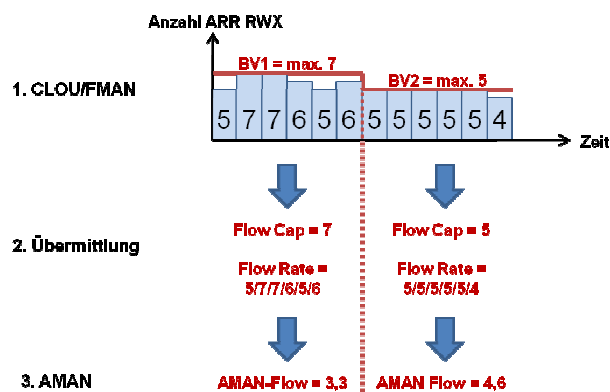


Abb. 3: Übermittlung CLOU/FMAN Flow an AMAN

Analog werden die optimierten Departure-Mengen je Zeitintervall an **darts** gesendet und von dem DMAN als Kapazität verarbeitet. Dabei wird nicht nur nach der Runway, sondern zusätzlich nach den jeweils für die Runway erlaubten Abflugfixen differenziert. So wird die Möglichkeit geschaffen, dem DMAN mitzuteilen, wenn ggf. Abflugströme von einer Startbahn auf eine andere verschoben werden. Dann erhält darts z.B. in Frankfurt für BIBTI Departures auf der RW25C die optimierte Menge 0 in den entsprechenden Zeitintervallen, da die Nordwestabflüge von CLOU/FMAN auf die RW18 geplant wurden. Zusätzlich ist eine Randbedingung in der Kopplung zu berücksichtigen: darts benötigt für jeden Abflug eine Runway-Zuweisung. Diese kann sich aufgrund der CLOU-Optimierung jedoch im Vergleich zum ursprünglichen Flugplan ändern. Somit wird für jeden Abflug die von CLOU festgelegte Runway mitgeschickt.

Das System ist weiter für die Bestimmung der TTOTs zuständig und berücksichtigt im Rechenprozess die Runway-Zuweisungen und die CLOU Flowwerte je Runway und je Abflugfixgruppe. Wenn darts für einen Flug keine Lösung findet, da für die Runway keine Kapazität für den Abflugfix des betroffenen Fluges zur Verfügung steht, muss das System in der Lage sein, die Departure-Runway auf die jeweils andere umzuschreiben. Anderenfalls würde der Flug solange in der Planung nach hinten geschoben werden, bis für den Abflugfix auf der gesetzten Runway wieder Kapazität vorhanden ist. Dies muss vermieden werden. Mit dem Senden der Zielzeit je Departure erfolgt auch eine Übermittlung der von darts umgesetzten Runway.

In der Gesamtübersicht ergibt sich damit folgender Systemverbund für die automatisierte Simulation (vgl. Abbildung 4):

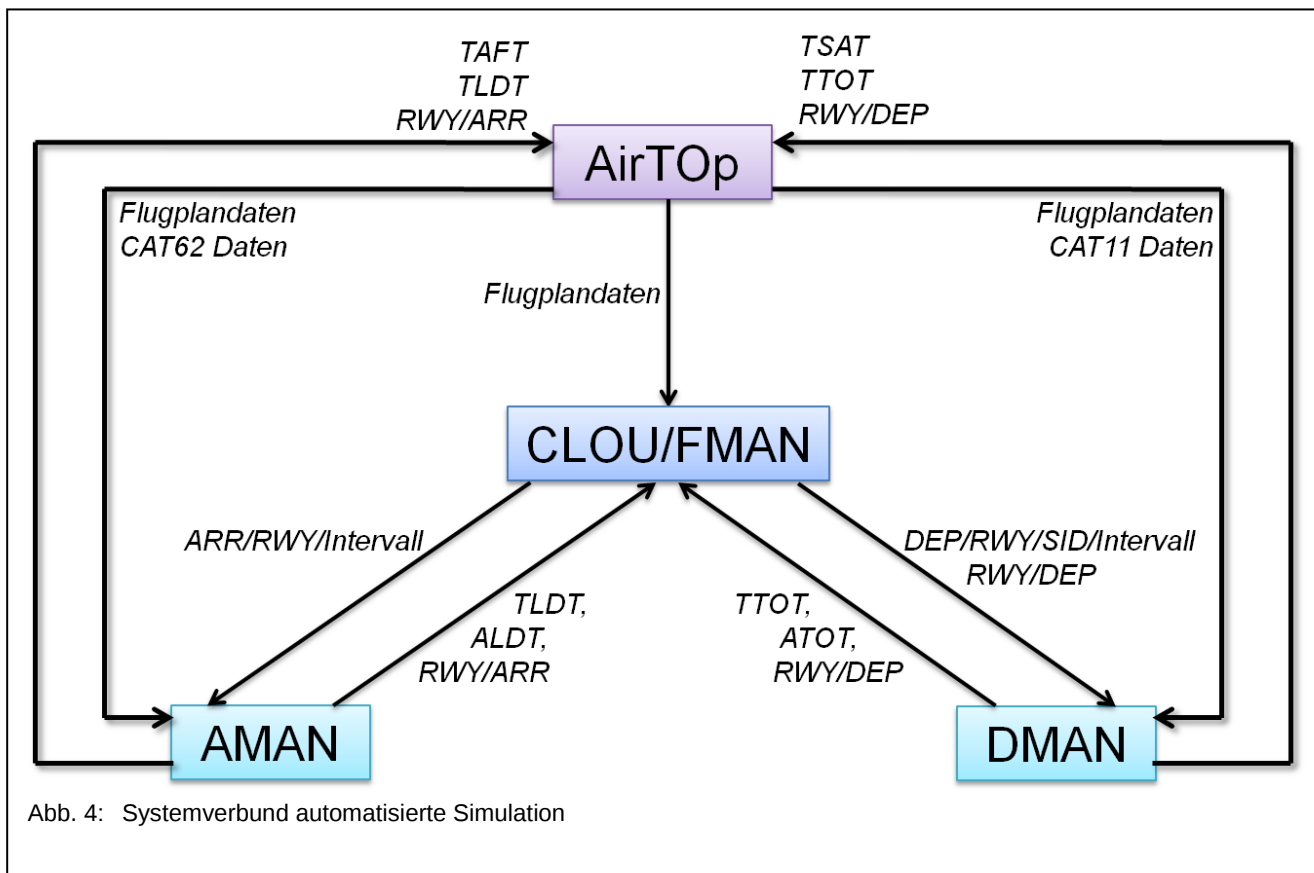


Abb. 4: Systemverbund automatisierte Simulation

Wie in Kapitel 2.1 erläutert, basiert die Berechnung des CLOU/FMAN auf einer Mengensteuerung unter Berücksichtigung definierter Teilkapazitäten für die einzelnen Verkehrsströme. Vorgaben von Radar- und Wirbelschleppenstaffelung bleiben unberücksichtigt. Die Sequenzplanungen in AMAN und DMAN funktionieren genau umgekehrt. Aus diesem Grund ist in dem Systemverbund sicherzustellen, dass CLOU/FMAN nicht größere Mengen an Arrivals und Departures je Runway vorplant, als die taktischen Systeme und letztendlich AirTop durch die Sequenzplanung umsetzen können. Die Feinjustierung muss vor Beginn der automatisierten Simulation durch Testläufe erfolgen, ggf. sind Teilkapazitäten im CLOU/FMAN anzupassen (vgl. Kapitel 6).

3.2 Erwartungen an die Planungsergebnisse

Zur Analyse des Nutzens einer prätaktischen Verkehrsvorsteuerung dient ein Vergleich der Verkehrsabwicklung ohne und mit CLOU/ FMAN Planung. Dafür ist eine genaue Definition des Referenz- und des Testszenarios sowie der einstellbaren Parameter erforderlich (vgl. Kapitel 6). Mit der Mengenvorsteuerung von CLOU/FMAN im Testszenario sollten sich aufeinander abgestimmte Arrival- und Departure-Ströme auf den Runways ergeben.

Im Referenzszenario kommt der CLOU/FMAN

nicht zum Einsatz, sodass keine Mengenvorsteuerung stattfindet und sich die Verkehrsströme in bestimmten Situationen gegenseitig negativ beeinflussen. So wird es für AirTop schwieriger, die Zielzeiten aus AMAN und darts umzusetzen, was zu Delay und Verspätungen führen kann.

Mit der CLOU/FMAN Planung kann erwartet werden, dass die erzeugten Delay-Minuten auf An- und Abflugseite insgesamt reduziert werden und der Durchsatz auf den Runways erhöht wird. Aus den primären Bewertungsparametern lassen sich weitere sekundäre wie Umwelteinflüsse ableiten. Eine genaue Beschreibung der Auswerteparameter ist in Kapitel 8 zu finden. Insgesamt kann die These aufgestellt werden, dass sich die Mengenvorsteuerung positiv auf die Verkehrsabwicklung sowie wirtschaftliche und ökologische Gesichtspunkte auswirkt.

Für die Auswertung des Nutzens einer prätaktischen Vorsteuerung sind Situationen interessant, in denen es aufgrund der Verkehrsnachfrage zu Engpässen kommt. Nur in solchen Fällen findet eine Optimierung statt. Dabei sind auch Effekte zu berücksichtigen, die sich durch einen Wechsel von einem Peak zu einem nächsten ergeben.

Folgende Tabelle 1 fasst die Erwartungen an die Simulationsergebnisse zusammen:

Validierungsziel	Hypothese	Indikator
optimierte Mengenvorsteuerung des ARR- und DEP-Verkehrs durch CLOU/FMAN	CLOU/FMAN optimiert die Bahnnutzungsstrategie in Engpasssituationen (Nachfrage erreicht Kapazitätsgrenzen)	Betriebsverfahrenoptimierung im CLOU/FMAN
	Der Delay wird im Vergleich zum Verkehrsablauf ohne CLOU/FMAN reduziert.	Delay (Vergleich Actual-Estimate)
	Die Verspätung wird im Vergleich zum Verkehrsablauf ohne CLOU/FMAN reduziert.	Pünktlichkeit (Vergleich Actual-Schedule)
	Der Durchsatz auf den Start- und Landebahnen wird im Vergleich zum Verkehrsablauf ohne CLOU/FMAN erhöht.	Durchsatz
	Die optimierte Mengen-Vorsteuerung durch CLOU/FMAN hat positive Effekte auf die wirtschaftliche und umweltgerechte Abwicklung des Verkehrs.	Treibstoffverbrauch CO ₂ -Ausstoß

Tabelle 1: Validierungsziele in der automatisierten Simulation

4. Rahmenbedingungen

4.1 Simulierter Flughafen und Luftraum

Die automatisierte Simulation wird auf Basis des Flughafens Frankfurt mit dem im Oktober 2011 in Betrieb gehenden 4-Bahn System durchgeführt, wobei lediglich die Betriebsrichtung 25 betrachtet wird.

Alle beteiligten Systeme werden soweit angepasst, dass sie in der Lage sind, den Flughafen abzubilden. Für CLOU/FMAN werden neue Bahnnutzungsstrategien aufgesetzt, die den Anforderungen des 4-Bahn Systems mit seinen komplexen betrieblichen Abhängigkeiten zwischen den Verkehrsströmen genügt. Dabei muss eine Balance zwischen der betrieblichen Umsetzbarkeit sowie ausreichend Optimierungspotential für den FMAN sichergestellt werden. Diese betrieblichen Abhängigkeiten sind auch von AirTop bei der Simulation der Verkehrsabwicklung zu berücksichtigen. Den implementierten An- und Abflugstrecken in AirTop sowie den taktischen Planungssystemen liegt der Entwicklungsstand von August 2011 für den Zeitpunkt der Inbetriebnahme der neuen Landebahn Nordwest im Oktober 2011 zugrunde.

4.2 Simulations- und Messgebiet

Das Messgebiet umfasst den Bereich, der für die Auswertungen herangezogen wird. In der automatisierten Simulation richtet sich das Messgebiet nach den Planungshorizonten der gekoppelten Systeme und schließt damit den Flughafen Frankfurt mit seinem Runway-Layout sowie den AMAN Planungshorizont mit einem Radius von 100 NM um jeden Metering Fix ein. Damit ist der Planungshorizont von darts mit abgedeckt. CLOU/FMAN betrachtet zwar einen größeren Zeitbereich, die Verkehrsbewegungen außerhalb des AMAN Horizontes müssen jedoch für die Durchführung der Simulation nicht simuliert werden. Für diese Schnellzeitsimulation ist es ausreichend, wenn CLOU/FMAN zunächst mit dem initialen Flugplan und anschließend den Updates arbeitet, die durch die Planungen der taktischen Systeme verursacht werden.

Zur realistischen Gestaltung der betrieblichen Ein- und Ausflugbedingungen wird um das Messgebiet herum das Simulationsgebiet erweitert.

5. Verkehrsbeispiel

Die neue Landebahn des Flughafens Frankfurt wird im Oktober 2011 in Betrieb gehen, sodass derzeit noch keine Live-Daten als Verkehrsbeispiel zur Verfügung stehen. Es wurden jedoch bereits verschiedene Verkehrsbeispiele mit unterschiedlichen Nachfrageszenarien für Voruntersuchungen im Rahmen von Schnellzeitsimulationen erstellt, die für die automatisierte Simulation verwendet werden können.

Dabei sind bestimmte Anforderungen, die aus dem Systemverbund resultieren, zu berücksichtigen. Die Nachfrage muss groß genug sein, um den Optimierungsprozess im CLOU/FMAN anzuregen. Gleichzeitig darf es jedoch nicht zu einer Überlastsituation am Flughafen einem Volllaufen des Flughafens kommen, um einen Deadlock in AirTop zu vermeiden, wenn dieser keine Lösung mehr berechnen kann. Bei einem Szenario für die komplette Ausbaustufe des Flughafens (126 Fb/h) besteht das Risiko, dass die getroffenen Annahmen für die Bahnnutzungsstrategien in der automatisierten Simulation nicht ausreichen, um den Verkehr ohne übermäßig großen Delay abzuwickeln. Aus diesem Grund wird von einem Szenario mit geringerer Nachfrage (d.h. nicht endgültige Ausbaustufe des Flughafens) ausgegangen. Des Weiteren wird keine Simulation der Bodenprozesse stattfinden. Zum einen ist dies technisch nicht erforderlich und eliminiert die Gefahr eines Deadlocks. Zum anderen mindert dies die Komplexität der automatisierten Simulation und hat keinen negativen Einfluss auf die Auswertung der grundsätzlichen Fragestellung. Die Simulation des Abflugverkehrs erfolgt über das Berechnen der TSAT und Ableiten der TTOT durch darts und des Erscheinens des Departures an der Schwelle zur TTOT, wobei das Luftfahrzeug ggf. durch ankommende Flieger verzögert wird.

Für eine andere Untersuchung im Rahmen einer Schnellzeitsimulation wurde der Prognoseflugplan TAAM77 aus dem Planfeststellungsbeschluss verwendet und von 126 Fb/h auf 100 Fb/h herunter gerechnet. Damit ergeben sich für den simulierten Tag (24 h) insgesamt 1683 Flugbewegungen, mit 839 Abflügen und 844 Anflügen. Dieses Verkehrsbeispiel wird als Grundlage auch für die Kopplung prätaktischer und taktischer Planungssysteme herangezogen, wobei das hinterlegte Routenlayout aktualisiert wird und

die Inhalte des Flugplans abgeändert werden. Um die übergreifenden Effekte der Verkehrssteuerung zwischen Arrival- und Departure-Peaks analysieren zu können, wird der gesamte Verkehrstag simuliert.

6. Simulationsszenarien

Im Folgenden werden das Referenzszenario und die Testszenarios beschrieben. Zuvor werden relevante Parameter festgelegt, um die Anzahl an durchzuführenden Simulationen einzuschränken.

6.1 Festlegung von Parametern

6.1.1 Mindestdauer eines Betriebsverfahrens

Die Mindestdauer eines Betriebsverfahrens im CLOU/FMAN ist grundsätzlich von der betrieblichen Anforderung abhängig. Sie muss eine gewisse Stabilität der Planungsvorschläge garantieren, aber auch ein gewisses Maß an Flexibilität ermöglichen. Für die automatisierte Simulation wird eine BV-Mindestdauer von 30 min festgelegt.

6.1.2 Optimierungsstufe

Wie unter Kapitel 2.1 erläutert, liefert der CLOU/FMAN drei Optimierungsstufen. Die sogenannte Natural-Lösung spiegelt dabei die heutige Arbeitsweise des Lotsen wider bei einer besseren Planungsgrundlage durch die Anzeige des Optimierungsvorschlags von CLOU. Die Flüge werden nach FCFS sortiert, zusätzlich wird jedoch bereits die optimierte Bahnnutzungsstrategie berücksichtigt. Die Flowopt-Lösung hält in den einzelnen Arrival- bzw. Departure-Strömen die FCFS-Sortierung aufrecht. Es können jedoch Verschiebungen zwischen Arrivals und Departures stattfinden, um die vorhandenen Kapazitäten noch besser auszunutzen. In der Target-Lösung findet keine Berücksichtigung der FCFS Sortierung mehr statt. Die Flüge können alle einzeln so gelegt werden, dass eine maximale Kapazitätsausnutzung möglich ist. Diese Lösung stellt in Hinblick auf die heutigen Arbeitsweisen des Lotsen eher eine theoretisch mögliche dar, die für die reale Verkehrsabwicklung jedoch uninteressant bleibt. Aus diesem Grund soll diese Optimierungsstufe in der automatisierten Simulation unberücksichtigt bleiben. Die

Flowopt-Lösung bietet im Vergleich zur Natural-Lösung ein größeres Optimierungspotential und wäre in der Realität umsetzbar. Deshalb kommt bei der Übermittlung der vorgesteuerten Verkehrsmengen an die taktischen Systeme diese Optimierungsstufe zum Einsatz.

Wenn die Simulationsergebnisse erkennen lassen, dass eine andere Einstellung dieser Parameter untersuchungswert ist, können weitere Analysen erfolgen.

6.1.3 Störgrößen

Im Rahmen der technischen Kopplungen zwischen AirTop und den Planungssystemen werden verschiedene Möglichkeiten implementiert, um manuelle Eingaben zu tätigen:

- manuelle Änderung der TOBT und TSAT
- manuelle Änderung der Startreihenfolge am Startbahnkopf
- manuelle Änderung von TAFT und TLDT
- manuelle Änderung von Geschwindigkeit, Höhe und Kurs bei Arrivals

Im Weiteren ist es möglich, im CLOU/FMAN manuell Betriebsverfahren und Teilkapazitäten zu ändern.

Bei der automatisierten Simulation mit insgesamt drei Planungssystemen kann jedoch davon ausgegangen werden, dass der Verkehrsablauf ohnehin dynamisch ablaufen wird. Mehrere Simulationen ein und des gleichen Szenarios können nicht zwangsläufig identisch ablaufen. Dafür sind die Planungsprozesse zu komplex und nichtdeterministisch. Aus diesem Grund ist die Generierung von Störgrößen über solche manuellen Eingaben und insbesondere die Analyse der Auswirkungen auf die Verkehrsabwicklung nur schwer nachvollziehbar. Als Alternative soll die AirTop-interne Funktionalität zur Generierung von stochastischen Störungen im Flugplan genutzt werden. Hierbei werden vor Beginn einer Simulation die Zeiten, zu denen die Flüge im Szenario auftauchen, mit Hilfe einer stochastischen Funktion verändert. Damit ändert sich auch der Demandverlauf über die Simulation hinweg.

6.2 Referenzszenario

In Hinblick auf das Untersuchungsziel, den Nutzen einer prätaktischen Vorsteuerung des Verkehrs zu analysieren, soll sich das Referenzszenario zum Testszenario lediglich in der Nichtbetrachtung der CLOU/FMAN Planungen unterscheiden, um zusätzliche Effekte auf die Verkehrsabwicklung von vornherein auszuschließen. Ein Departure Manager im Sinne von darts existiert derzeit nicht im operativen Betrieb in Frankfurt, für die Kopplung wird er aus genanntem Grund dennoch herangezogen. D.h., im Referenzszenario wird mit den Planungen von AMAN und DMAN gearbeitet. In diesem Fall gibt es keine automatische Steuerung des AMAN-Flows. Ein zu Beginn eingestellter und über den Simulationstag hinweg statischer Flow würde die Ergebnisse des Referenzszenarios jedoch von vornherein verschlechtern, was zu einer Verzerrung der Mehrwertanalyse führt. Auch ohne eine CLOU/FMAN Vorsteuerung muss eine Anpassung des AMAN-Flows an die Nachfrage sichergestellt sein, da dies im operativen Betrieb eine Standardprozedur ist. Mit dem bekannten Demandverlauf des Flugplans werden Zeitintervalle für Departure- und Arrival-Peaks definiert und entsprechende AMAN-Flowwerte hinterlegt. Diese sind vorher durch Vergleiche aus dem derzeitigen 3-Bahn Betrieb von Demandsituation und eingestelltem AMAN-Flow unter Berücksichtigung der Wettersituation auszuarbeiten.

Alle weiteren eingestellten Parameter sind identisch zum anschließenden Testszenario. Es wird mit dem Szenarioflugplan mit einem Eckwert von 100 Fb/h gerechnet.

Folgende Abbildung 5 veranschaulicht den planerischen Ablauf des Referenzszenarios:

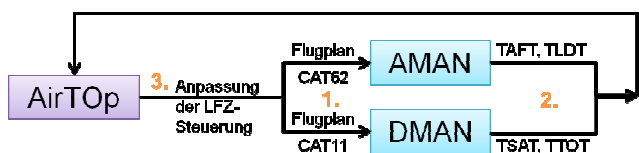


Abb. 5: Planerischer Ablauf Referenzszenario

Das Referenzszenario wird zum einen ohne und zum anderen mit Störgrößeneinflüssen simuliert.

6.3 Testszenario

Im Testszenario kommt zusätzlich zum Referenzszenario der CLOU/FMAN zum Einsatz, womit sich der Planungsablauf im Systemverbund im Vergleich zu Abbildung 5 verändert (Abbildung 6):

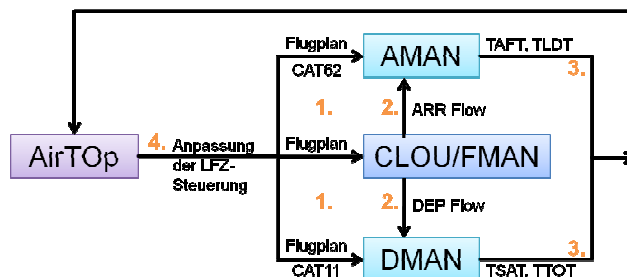


Abb. 6: Planerischer Ablauf Testszenario

Es wird mit demselben Flugplan wie im Referenzszenario gerechnet. Die Kapazität, die im CLOU/FMAN als Planungsgrundlage hinterlegt werden muss, entspricht dem Eckwert von 100 Fb/h.

Auch das Testszenario wird einmal ohne und einmal mit Störgrößeneinflüssen simuliert.

6.4 Anzahl an Simulationen

Es ergeben sich zunächst vier Simulationsszenarien:

- (1) Referenzszenario I: AMAN Planung, darts Planung, keine Störgrößeneinflüsse
- (2) Referenzszenario II: AMAN Planung, darts Planung, mit Störgrößeneinflüssen
- (3) Testszenario I: CLOU/FMAN Planung, AMAN Planung, darts Planung, BV-Mindestdauer 30 min; Flowopt-Lösung, keine Störgrößeneinflüsse
- (4) Testszenario II: CLOU/FMAN Planung, AMAN Planung, darts Planung, BV-Mindestdauer 30 min; Flowopt-Lösung, mit Störgrößeneinflüssen

7. Durchführung der automatisierten Simulation

Vor der Durchführung der eigentlichen automatisierten Simulation sind mehrere Testläufe mit dem etablierten Systemverbund einzuplanen, um das Zusammenspiel der Einzelsysteme zu testen

und ggf. weitere erforderliche Anpassungen vorzunehmen. Voraussetzung für die automatisierte Simulation ist neben der technischen Kopplung der Systeme eine stabile Planung des Verkehrsablaufs. Des Weiteren muss geprüft werden, ob die eingestellten Parameter verändert werden müssen, sodass eine sinnvolle und nachvollziehbare Planung zu erkennen ist.

einen Zeitpuffer für eventuell notwendige letzte Anpassungen vorzuhalten.

Folgende Tabelle 2 fasst zusammen, welche Punkte in den Testläufen zu betrachten sind:

Untersuchungspunkt	Hinweise
Stabilität der AMAN Sequenzplanung	Stabilität kann beeinträchtigt werden durch ➤ CLOU/FMAN Flow-Vorgaben ➤ Neusteuerung der Arrivals in AirTop, der TAFT und TLDT versucht umzusetzen
Stabilität der darts Sequenzplanung	Stabilität kann beeinträchtigt werden durch ➤ CLOU/FMAN Flow-Vorgaben ➤ Neusteuerung der Departures in AirTop, der TTOT versucht umzusetzen
Umsetzung der taktischen Zielzeiten in AirTop	Umsetzung wird nicht exakt möglich sein; es darf jedoch nicht zu einer permanenten Rückkopplung der Neusteuerung auf die Sequenzplanung kommen
CLOU/FMAN führt Optimierungsprozess vollständig durch	unkritisch, solange AMAN und darts die CLOU/FMAN Flow-Vorgaben exakt einhalten; wenn mehr Luftfahrzeuge im CLOU/FMAN fixiert werden, als Kapazität im Intervall zur Verfügung steht, wird Optimierungslauf abgebrochen
alle Flüge werden in AirTop bis zur ATOT bzw. ALDT simuliert	in AirTop ist es möglich, dass einzelne Arrivals oder Departures während des Simulationslaufs gelöscht werden; das darf nicht vorkommen
Teilkapazitäten im FMAN	ggf. anpassen, damit die Anzahl der möglichen Flugbewegungen mit der Steuerung in AirTop übereinstimmen
Planungshorizont, in dem Zielzeiten fixiert werden	derzeit auf 30 min gesetzt; Auswirkungen einer Änderung prüfen und ggf. anpassen
Länge der Intervalle für CLOU-Flowübermittlung	derzeit auf 10 min gesetzt; Auswirkungen einer Änderung prüfen und ggf. anpassen

Tabelle 2: Untersuchungspunkte für Testläufe mit dem Systemverbund

Nach Abschluss der Testläufe ist in einer Generalprobe sicherzustellen, dass der Systemverbund in der automatisierten Simulation ohne Komplikationen läuft und die Planungen korrekt durchgeführt werden. Die Generalprobe soll mindestens zwei Wochen vor Beginn der eigentlichen automatisierten Simulation stattfinden, um

In der automatisierten Simulation werden die in Kapitel 6 definierten Referenz- und Testszenarien nacheinander simuliert und im Anschluss die Ergebnisse ausgewertet.

Nach derzeitigem Planungsstand wird die Durchführung der automatisierten Simulation im März 2012 stattfinden. Die Testläufe haben im Sep-

tember 2011 begonnen und sollen bis Mitte Februar 2012 abgeschlossen sein. Die Auswertung der Ergebnisse wird voraussichtlich bis Ende Mai 2012 fertiggestellt sein.

8. Auswertungskriterien

Tabelle 1 aus Kapitel 3.2 gibt bereits eine Übersicht über die relevanten Auswerteparameter, um den Mehrwert der prätaktischen Vorsteuerung zu bewerten.

- Die Abfolge der optimierten Betriebsverfahren ist ein qualitatives Indiz dafür, inwieweit der CLOU/FMAN auf Departure- und Arrival-Peaks reagiert.
- Über die Höhe des im Szenario produzierten Delays sowie der Verspätung (Maß für Pünktlichkeit) kann quantitativ erfasst werden, welchen positiven Effekt die Mengenvorsteuerung auf die Verkehrsabwicklung hat. In AirTOP lässt sich diesbezüglich der sogenannte Sequencing Delay errechnen. Hierzu wird als Referenz zur ATOT bzw. ALDT die erwartete Lande- bzw. Startzeit herangezogen, die erreicht werden würde, wenn das Luftfahrzeug das System ohne Störung auf dem kürzesten möglichen Weg passieren würde. Eine Unterscheidung zwischen einer Schedule Time (als Referenz für Pünktlichkeit) und einer Estimate Time (als Referenz für Delay) ist nicht möglich.
- Der Durchsatz je Start- und Landebahn gibt ebenfalls eine quantitative Aussage über den potentiellen Nutzen der Mengenvorsteuerung.

Die Parameter werden über den gesamten Simulationszeitraum hinweg erfasst. Der Verlauf von Delay und Durchsatz wird in rollierender Stunde dargestellt, wobei die Werte mit einem Abstand von 10 min präsentiert werden.

Für die Auswertung sind besonders die Zeiträume relevant, in denen eine Optimierung des CLOU/FMAN stattgefunden hat. Auf diese soll der Fokus gerichtet werden.

Zur zusätzlichen Bewertung der Ergebnisse werden folgende Umweltparameter abgeleitet:

- Kerosin-Verbrauch
- CO₂-Ausstoß

Hierfür wird ein Auswertetool verwendet, das im iPort AP1.4.2 (Auswertemöglichkeit Emissionen) entwickelt wird. Dieses Tool liefert im Ergebnis die Menge an Treibstoff in Tonnen sowie die davon abgeleitete Menge an CO₂ in Tonnen, die über den gesamten Simulationszeitraum hinweg produziert wird. Das Tool ist in der Lage, die von AirTop ausgegebenen Ergebnisdaten einzulesen und die Berechnungen durchzuführen. Zum Zeitpunkt der Fertigstellung dieses Konzepts gibt das Tool die Summen über die genannten Parameter aus. Zum Zeitpunkt der Auswertung der durchgeführten Simulation sind ggf. weitere Möglichkeiten nutzbar, wie die Ermittlung des Treibstoffverbrauchs je Luftfahrzeug über den simulierten Tag hinweg.

9. Perspektiven

Nach Abschluss dieses iPort Teilprojekts kann eine Aussage dazu getroffen werden, inwieweit eine prätaktische Mengenvorsteuerung einen Nutzen für die Verkehrsabwicklung an einem Hub-Flughafen bringt (vgl. Kapitel 3). Daraus können sich ggf. weitere Projekte zur Erarbeitung neuer Ansätze in der unterstützenden Systemlandschaft des operativen Betriebs ergeben.

Daneben findet im Rahmen dieses Projekts zum ersten Mal eine Kopplung zwischen einem Schnellzeitsimulationstool und einem bzw. mehreren Planungssystemen statt. Mit der automatisierten Simulation kann gezeigt werden, dass solch eine Systemkopplung technisch möglich ist und für Untersuchungszwecke sinnvoll genutzt werden kann. Diese Art von Kopplung könnte in Zukunft für weitere Projekte genutzt werden. Insbesondere für die Erprobung neuer Funktionen eines Planungssystems kann solch eine Kopplung von Nutzen sein, um den Aufwand für Vorbereitung und Durchführung von Realzeitsimulationen zu reduzieren.

Nach Abschluss dieses Projekts wird kommenden Jahr ein weiterer TE im Fokus Artikel folgen, der die Ergebnisse und den weiteren Ausblick vorstellt.

Abkürzungen

AMAN	Arrival Manager
ALDT	Actual Landing Time
ATOT	Actual Take-off Time
BV	Betriebsverfahren
CLOU	Cooperative Local Resource Planner
DMAN	Departure Manager
FCFS	First Come First Serve
FMAN	Flow Manager
iPort	innovativer Airport
K-ATM	Kooperatives Air Traffic Management
TAFT	Target Initial Approach Fix Time
TLDT	Target Landing Time
TOBT	Target Off-block Time
TSAT	Target Startup Approval Time
TTOT	Target Take-off Time
WFF	Wettbewerbsfähiger Flughafen

Referenzen

[1]	E. Rehwald, P.Hecker (TU BS), S. Marx, K. Nachtigall (TU DD), K. Schwentek (DFS): A Prototype of an pre-tactical Airport Centered Flow Management, veröffentlicht auf: EIWAC 2010, 12.11.2010
[2]	Katharina Schwentek,E. Rehwald, P. Hecker (TU Braunschweig) Bahnnutzungsstrategien bei einem Flughafen-zentrierten Flowmanagement , Deutscher Luft- und Raumfahrtkongress 2010, 01.09.2010
[3]	Bergner, J.; König, C.; Hofmann, T.; Ebert, H. : Entwurf einer integrierten Planungsanzeige für den Tower-Controller, Deutscher Luft- und Raumfahrtkongress 2008
[4]	Dr. Rainer Kaufhold, Dr. Jörg Bergner: K-ATM Abschluss, Teilvorhaben KOPIM-ATC: Erweiterung AMAN Planungshorizont, 28.11.2007
[5]	Dr. Rainer Kaufhold: Abschluss des Vorhabens "Kooperatives Air Traffic Management", TE im Fokus 01/2008
[6]	Andreas Türk, Oliver Rühl: Der neue iPort-Experimentalplatz , TE im Fokus 02/2010

Kalibrierung und Validierung eines taskbasierten Arbeitslastmodells für Flugsicherungssysteme der DFS

Dr. Ralph Leemüller, Stefan Tenoort (TE/S), Theresa Emsbach (FH Aachen)

Zusammenfassung¹

Das bisher genutzte Arbeitslastmodell für Flugsicherungssysteme der DFS wurde durch ein taskbasiertes Arbeitslastmodell zur Ermittlung von Lotsen-Belastungen und Sektor-Kapazitäten ersetzt. Der vorliegende Beitrag beschreibt das entwickelte Workload-Modell. Darüber hinaus wird der Prozess der Validierung der Workload-Parameter für die Niederlassung Karlsruhe mit der seit 2011 zur Verfügung stehenden P1/VAFORIT-Umgebung dargestellt. Das taskbasierte Workload-Modell erbringt auf der Basis der neuen Task-Parameter plausible Ergebnisse für die Kriterien Belastung und Kapazität und bildet die Belastung durch das genutzte Flugsicherungs- (Air Traffic Control-, ATC-)System adäquat ab. Eine Verschiebung der Beanspruchung innerhalb des Teams vom Radar- zum Planungsloten wurde anhand der Kalibrierung bestätigt. Das entwickelte Modell ist eine Grundlage für die Berechnung von Kapazitätswerten für die Sektoren der Niederlassung Karlsruhe mit dem System CAPAN (Capacity Analyser).

1 Einleitung

Aufgrund des stetig anwachsenden Flugverkehrs nehmen die Anforderungen an das Flugverkehrsmanagement kontinuierlich zu. Um mit den steigenden Zahlen des Flugaufkommens Schritt halten zu können und den damit einhergehenden Ansprüchen gerecht zu werden, wie beispielsweise den Flugverkehr auch zukünftig sicher kontrollieren zu können, ist eine kontinuierliche Weiterentwicklung der Technologien und Systeme unumgänglich. Diese müssen es dem Lotsen ermöglichen, den gesteigerten Flugverkehr bei mindestens gleich bleibendem Sicherheitsstandard zu bewältigen.

¹ Dieser Artikel ist in Teilen veröffentlicht in: „Ergonomie im interdisziplinären Gestaltungsprozess“, DGLR Bericht 2011-1, Hrsg.: M. Grandt, S. Schmerwitz, Bonn 2011.

Im Bereich des Flugverkehrsmanagements ist vor allem das Berufsfeld des Lotsen stark von den fortschreitenden technologischen Entwicklungen geprägt. Denn Lotsen agieren innerhalb eines komplexen und dynamischen sozio-technischen Umfeldes, welches grundsätzlich einen hohen Grad an Multitasking-Fähigkeiten erfordert. Die Vielschichtigkeit der benötigten Fähigkeiten und besonderen mentalen Anforderungen dieser Berufsgruppe reicht von überdurchschnittlichem Konzentrationsvermögen, ausgeprägtem räumlichen Vorstellungsvermögen und gutem Gedächtnis, bis hin zu hohen Kommunikations- und Teamfähigkeiten und die Fähigkeit Entscheidungen unter Zeitdruck treffen zu können (Salas, 2010). Zu diesen natürlichen Belastungsquellen kommen zusätzliche Beanspruchungen aufgrund des kontinuierlich wachsenden Verkehrsstromes und der damit einhergehenden Technisierung der Systeme hinzu.

2 Bestimmung von Arbeitsbeanspruchung

Betrachtet man den Stellenwert, welchen die Arbeitsbeanspruchung aufgrund des Einflusses auf die Sicherheit einnimmt, ist die Vielzahl der durchgeführten Studien und damit der unterschiedlichen Ansätze zur Messung der Beanspruchung und für die Entwicklung von Workload-Modellen leicht verständlich. Im Folgenden soll auf die für diese Studie relevanten Messverfahren bzw. Workloadmodelle eingegangen werden.

In Abbildung 1 ist eine Übersicht der gängigen Messverfahren zur Erfassung der Arbeitsbeanspruchung dargestellt (Salas, 2010). Prinzipiell lässt sich sagen, dass keine optimale Methode zur Messung der Arbeitsbeanspruchung eines Lotsen existiert. Alle diese Verfahren besitzen Vor- und Nachteile, welche im Falle einer Anwendung zu beachten sind, um so eine gewissenhafte Analyse und deren Interpretation zu gewährleisten. Es bietet sich daher an, sich nicht nur auf eine einzige dieser Methoden zu be-

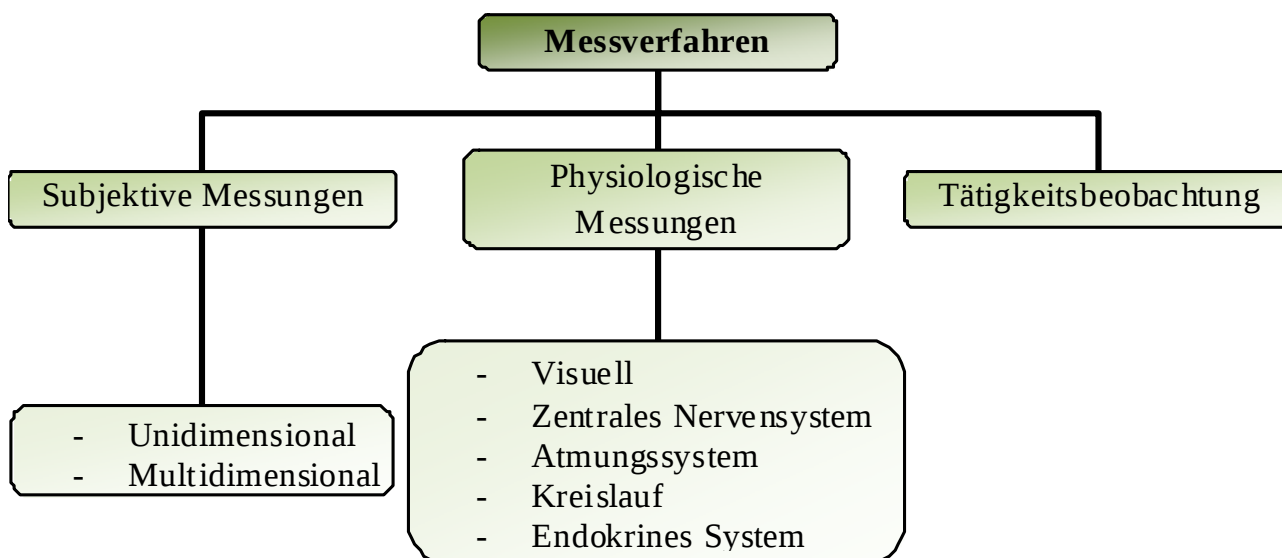


Abbildung 1: Übersicht der Methoden zur Erfassung der Arbeitsbeanspruchung

schränken, sondern auf eine methodenübergreifende Kombination aus diesen Verfahren zurückzugreifen. Dadurch können die Vor- und Nachteile ausgewählter Methoden gegeneinander abgewogen werden.

2.1 Subjektive Erhebungen

Zur Bestimmung der Arbeitsbeanspruchung eines Lotsen wird häufig auf subjektive Messverfahren zurückgegriffen. Im Rahmen dieser Methode geben die befragten Lotsen eine eigene Einschätzung über Art und Niveau der erlebten Arbeitsbeanspruchung ab. Um diese Art der Datenerhebung einheitlich durchführen zu können, werden vorab festgelegte Maßeinteilungen genutzt, sogenannte Rating-Skalen. Diese können ein- oder mehrdimensional sein. Die Dimension der Rating-Skalen beschreibt die Anzahl der betrachteten Faktoren. Ist ein Modell multi- bzw. mehrdimensional, so schließt es mehrere Parameter in die Analyse ein, wo hingegen ein- oder unidimensionale Ansätze den Fokus lediglich auf einen Faktor legen.

Subjektive Methoden finden häufig Anwendung zum einen, weil sich die Durchführung der Messung verhältnismäßig unkompliziert und vor allem preiswert arrangieren lässt, zum anderen jedoch aufgrund der hohen Empfindlichkeit gegenüber den Arbeitsanforderungen vor allem in Bereichen unterhalb der Leistungsfähigkeitsgrenze.

2.2 Bestimmung der Arbeitsbeanspruchung anhand von Situationsparametern

Eine Möglichkeit die Arbeitsbeanspruchung der Lotsen anhand von Workloadmodellen zu bestimmen, ist die Betrachtung von Situationsparametern. Diese Art der Beanspruchungsbewertung wird in der Literatur auch als eventbasierte Methode bezeichnet. Als Situationsparameter gelten unter anderem Faktoren wie z.B. Anzahl und Aufenthaltsdauer der Luftfahrzeuge im Sektor oder Vertikalbewegung. Die Abschätzung der Arbeitsbeanspruchung der Lotsen erfolgt anhand dieser Situationsparameter (Udovic & Tenoort, 2001).

Gerade dieser Ansatz zur Abschätzung der Arbeitsbeanspruchung ermöglicht eine Vielzahl von verschiedenen Herangehensweisen. Unterschiede der einzelnen Modelle sind vor allem in Anzahl und Art der involvierten Parameter zu erkennen.

Eine wesentliche Schwäche dieser Modelle ist die fehlende oder nur eingeschränkt mögliche Berücksichtigung der zugrundeliegenden technischen Systeme, d.h. der eingesetzten Flugsicherungssysteme, verfügbarer Unterstützungssysteme sowie Arbeitsverfahren. Aufgrund der vorliegenden Berechnungsvorschriften bliebe die mit einem neuen Flugsicherungssystem verbundene Arbeitsbeanspruchung unter Zugrundelegung identischer Verkehrssituationen gegenüber einem alten System also grundsätzlich unverändert. Eine weitere Schwäche ist die Verwendung

von Arbeitslastpunkten, die sich einer anschaulichen Beschreibung durch die Beiträge der einzelnen Parameter entziehen; dies ist z.B. bei der festgelegten Arbeitslastgrenze von 90 Punkten erkennbar. Ebenso ist es unklar, ob eine Steigerung von 60 auf 90 Arbeitslastpunkten auch einer Erhöhung der subjektiven Arbeitsbeanspruchung um 50% entspricht.

2.3 Arbeitszeitbasierter Ansatz zur Bestimmung der Arbeitsbeanspruchung

Grundlage aller Modelle, die auf einem arbeitszeitbasierten oder auch taskbasierten Ansatz beruhen, ist die Identifikation der Lotsentätigkeiten mit Hilfe einer Tätigkeitsanalyse und Annahme eines Zeitbudgets für jede dieser Tätigkeiten (Udovic & Tenoort, 2001).

Um diese Abschätzung der Arbeitsbeanspruchung durchführen zu können, müssen zunächst die Tätigkeiten der Lotsen genauer betrachtet werden. Dabei ist lediglich ein Teil der Aktivitäten direkt sichtbar, die kognitiven/mental Anteile entziehen sich direkten Beobachtungen. Eine Erfassung der sichtbaren Tätigkeiten kann über Beobachtungen erfolgen. Unterschiede weisen die Modelle vor allem bei der Einordnung der Tätigkeiten und damit im Detaillierungsgrad ihrer Erfassung auf. Jeder Lotsentätigkeit wird anschließend eine genaue, normative Bearbeitungszeit zugewiesen. Diese kann durch Beobachtung empirisch erhoben oder auch von Experten abgeschätzt werden.

Zur Ermittlung der Arbeitsbeanspruchung werden die durchgeführten Tätigkeiten des Lotsen über einen festgelegten Zeitraum, meist eine Stunde, aufgenommen und anschließend mit den zuvor determinierten Arbeitszeiten verrechnet und aufsummiert. Die dadurch ermittelte Arbeitszeit wird in Relation zum Zeitraum der Messung gesetzt. Demnach ist die errechnete Arbeitszeit identisch mit der Zeit, die der Lotse am Radarschirm verbringt, wenn die Arbeitsbeanspruchung des Lotsen 100% beträgt. Da es allerdings wenig erstrebenswert ist, den Lotsen ständig an seiner maximalen Leistungsfähigkeit arbeiten zu lassen, wird ein Grenzwert der maximalen Dauerbeanspruchung festgelegt. Sobald die aufsummierte Arbeitszeit 70% des vorgegebenen Zeitintervalls, das entspricht 42 Minuten bei einer einstündigen Messdauer, erreicht oder gar übersteigt, ist auch

die Kapazitätsgrenze erreicht. Diese Definition der Grenze wurde in mehreren Studien bestätigt. Es wird damit auch ein Sicherheitspuffer für mentale Aktivitäten eingeräumt, die nicht einer Beobachtung zugänglich sind, bzw. nicht im Arbeitslastmodell verlässlich abgebildet sein könnten.

Bei Erstanwendung dieser Modelle ist jedoch ein größerer Aufwand notwendig, um die Tätigkeiten eines Lotsen im Vorfeld anhand von Tätigkeitsanalysen zu identifizieren. Diese müssen aufgrund veränderter Arbeitsprozesse gegebenenfalls bei Implementierung neuer Systeme wiederholt werden. Auch voneinander abweichende Systemlandschaften und Verfahren der einzelnen Niederlassungen können zusätzliche Untersuchungen erfordern, um die Task-Parameter durch Ermittlung neuer Tätigkeiten und deren Arbeitszeiten an veränderte Bedingungen, wie z.B. neue Systeme, Tools oder Arbeitsweisen, anzupassen. Wie bereits beschrieben, stellt sich auch hier die Schwierigkeit, nicht beobachtbare Tätigkeiten oder parallel ablaufende kognitive Prozesse zu bestimmen.

3 Neues Workload-Modell aufgrund ATC-Systemwechsel bei der DFS

Das neue ATC-System P1/VAFORIT der DFS ist 2010 in Betrieb gegangen und hat das mainframebasierte System KARLDAP ersetzt, das mehr als 30 Jahre lang in Betrieb war. Das System ist im Center Karlsruhe für die Streckenkontrolle im oberen Luftraum im Einsatz. Die vierdimensionale Trajektorienvorhersage sowie die Abkehr von Kontrollstreifen hin zum kontrollstreifenlosen Betrieb stehen für einen Paradigmenwechsel in der Flugdatenverarbeitung. Solche trajektorienbasierten Systeme stehen auch im Zentrum des Forschungsprogramms für das Flugverkehrsmanagement im einheitlichen europäischen Luftraum (Single European Sky, SES). Sie stellen vierdimensionale Daten für die geplanten Flugwege aller für die Kontrollzentrale relevanten Flüge bereit. Diese Funktionen ermöglichen es den Lotsen, mögliche Konflikte zwischen Luftfahrzeugen frühzeitig zu erkennen und zu vermeiden. Außerdem erhöhen sie die Genauigkeit der Flugplanung, so dass die Pünktlichkeit verbessert und Umwege reduziert werden können.

Die wichtigsten Funktionen von P1/VAFORIT sind:

- 4D-Trajektorienvorhersage
- Hochentwickelte Tools für Lotsen, wie Medium Term Conflict Detection (MTCD), What-if Probing, Executive Conflict Search (ECS) sowie Flight Conformance Monitoring Aids (MONA)
- Kontrollstreifenloser Betrieb
- Short Term Conflict Alert (STCA) und Area Proximity Warning (APW)
- Automatische Center-Center-Koordination (OLDI 4.1)
- Mode-S-Verarbeitung

Zur Analyse der Kapazität für Flugsicherungssysteme wird das Tool CAPAN eingesetzt. CAPAN basiert auf einem eventbasierten Workload-Modell, mit dem die Grenzen der Beanspruchung für jeden einzelnen Sektor berechnet werden. Dieses Modell lieferte bisher sehr gute Ergebnisse bei streifenorientierten ATC-Systemen. Für das neue trajektorienbasierte System kann das Kapazitätsanalysetool CAPAN allerdings keine adäquaten Aussagen liefern. Daher ist die Entwicklung eines taskbasierten Arbeitslastmodells erforderlich.

Das Workload-Modell wird auch zur Entscheidungsunterstützung bei der Entwicklung und Erprobung moderner Luftraumstrukturen im Rahmen von Schnellzeitsimulationen eingesetzt. Dabei ist es erforderlich, das Potential der Strukturen zu bewerten und die Auswirkungen auf die Belastung der Fluglotsen zu erfassen. Die Bestimmung der Arbeitslast der Lotsen soll zur Stärkung der Akzeptanz mit einem übersichtlichen und transparenten Modell erfolgen und muss Änderungen der Belastungen durch veränderte Luftraumstrukturen, Arbeitsmethoden und neue Unterstützungssysteme abbilden. Dabei ist es unerlässlich, Aussagen hinsichtlich des Potentials neuer Strukturen und der Grenzbelastung zu treffen. Eine niedrigere Belastung eröffnet die Möglichkeit, zusätzliche Flugbewegungen zu kontrollieren, und erlaubt damit eine höhere Luftraumkapazität.

Es ist für die Schnellzeitsimulation nicht notwendig, ein Arbeitslastmodell einzusetzen, das die Arbeitsprozesse der Lotsen so detailliert wie möglich abbildet. Vielmehr muss gewährleistet sein, dass die Arbeitsbelastung

der Lotsen entsprechend der jeweiligen Verkehrssituation hinreichend genau abgebildet werden kann. Das Arbeitslastmodell in der Schnellzeitsimulation und CAPAN muss dasselbe sein, da die Ergebnisse beider Systeme vergleichbar sein sollen.

Bei der Entwicklung des taskbasierten Arbeitslastmodells sollen folgende Kriterien beachtet werden:

- Bestimmung der Arbeitsbelastung des Lotsen in den Bereichen Area Control Center (ACC, Streckenkontrolle unterer Luftraum), Upper Area Control Center (UAC, Streckenkontrolle oberer Luftraum) und Approach (APP, Anflugkontrolle),
- Determiniertheit bzw. Wiederholbarkeit der Simulation, d.h. keine Beeinflussung des Simulationsverlaufs durch das Arbeitslastmodell,
- hohe Transparenz durch die Nutzung einer überschaubaren Anzahl von Tasks,
- grundsätzliche Übereinstimmung der Ergebnisse mit denen aus bewährten Simulationsmodellen.

Bei der Entwicklung des Arbeitslastmodells wurde auf die Ideen existierender Workload-Modelle zurückgegriffen. So werden teilweise Elemente verwendet, die ursprünglich von Eurocontrol für den dort entwickelten RAMS-Simulator genutzt wurden. Andere Tasks lehnen sich an Verfahren an, die dem existierenden Arbeitslastmodell des Forschungsbereichs der DFS entnommen wurden.

3.1 Beschreibung der Tasks

Das taskbasierte Arbeitslastmodell besteht aus sechs Task-Gruppen. Es beinhaltet bezüglich der Aufgaben und Taskzeiten eine Unterscheidung für den oberen und unteren Luftraum sowie den Anflugbereich. Aus der Lotsenfunktion ergibt sich eine weitere Differenzierung für Radar- und Planungsloten (Abbildung 2).



Abbildung 2: Task-Bereiche der Lotsen

3.2 Monitoring

Das Monitoring umfasst die Grundbelastung bzw. Routinetätigkeiten eines Radar- und Planungsloten.

Zur Berechnung der Teilarbeitslast werden als Eingangsgröße die Anzahl der Luftfahrzeuge und ihre Verweildauer im Sektor herangezogen. Jedes Luftfahrzeug wird bei Einflug in den Sektor erstmals erfasst. In bestimmten Abständen wird für jedes Luftfahrzeug die Überwachungsaufgabe („Monitoring“) wiederholt, um beispielsweise Kontrollaufgaben in Bezug auf das Einhalten der Flughöhe und des Flugweges etc. zu erfassen.

Der Planungslotse unterstützt den Radarlotsen bei der Überwachung der Luftfahrzeuge, führt diese Aufgabe aber nicht so intensiv aus wie der Radarlotse. Die Taskdauer für das Monitoring und das Beobachtungsintervall werden daher für den Planungslotse geringer angesetzt.

3.3 Radio Telephony (RT)

Diese Task bildet die Belastung des Radarlotsen durch Routinefunk ab. Hierbei wird zwischen dem Initial Call, also dem „Erstanruf“, und dem „Transfer Call“ also dem „Wegschicken“ in den nächsten Sektor unterschieden.

3.4 Freigaben (Clearance)

Will ein Luftfahrzeug seine Reiseflughöhe verlassen, um den Landeanflug zu beginnen, oder muss ein Luftfahrzeug aufgrund einer Kurs- bzw. Routenänderung eine andere Flugfläche einnehmen, so muss der Radarlotse dem Luftfahrzeugführer eine Freigabe für den Steig- bzw. Sinkflug erteilen.

3.5 Konfliktsuche

Mit der Task Konfliktsuche wird die Arbeitszeit des Radar- und Planungsloten erfasst, die benötigt wird, um seinen Luftraum nach potentiellen Konflikten zu durchsuchen und deren Lösungen zu planen. Ein potentieller Konflikt ist eine Annäherung zweier Luftfahrzeuge bei Unterschreitung der Mindestradarstaffelung. Da in der Phase der Konflikterkennung, einige Minuten vor einem sich anbahnenden Konfliktereignis, die Prognose für die Annäherung relativ grob erfolgt, werden in der Simulation auch Annäherungen zwischen

100% bis 300% des Mindestradarstaffelungswertes erfasst und für die Errechnung der Belastung verwendet.

Der Arbeitsaufwand, den ein Konflikt verursacht, hängt auch von der Anzahl der Luftfahrzeuge ab, die sich bereits im Sektor befinden (die sogenannte Occupancy). Werden vom Lotsen viele Luftfahrzeuge im Sektor geführt, steigt die Konfliktsuche (Analyse der entsprechenden Trajektorien) aufgrund der Anzahl der möglich beteiligten Luftfahrzeuge überproportional an. Dies wird durch eine Erweiterung des Arbeitslastmodells berücksichtigt. Die dafür hinterlegte Zeit wird nun für jeden Konflikt in Abhängigkeit von der Occupancy mit einem zusätzlichen Faktor multipliziert.

Eine Zunahme der Workload des Tasks Konfliktsuche bei hoher Occupancy kann damit gut dargestellt werden, und Arbeitslastspitzen werden stärker betont. Abbildung 3 zeigt die Conflict Detection Workload für einen Sektor mit und ohne Berücksichtigung der Occupancy. In der Spitze steigt die Arbeitslast für die Konflikterkennung um 1,5 Minuten oder etwa 30 Prozent.

3.6 Konfliktlösung

Die Aufgabe der Konfliktlösung wird von Radar- und Planungsloten durchgeführt. Diese Aufgabe umfasst die Umsetzung einer Lösung in Form von Anweisungen an den Piloten zur Vermeidung von Konfliktsituationen.

3.7 Koordination

Da moderne Flugsicherungssysteme automatisch Ein- und Ausflugszeiten koordinieren und weitgehend standardisierte Übergabehöhen benutzen, ist der telefonische und elektronische Koordinationsaufwand gering. Dennoch müssen einzelne vom Zeitplan, Flugweg und Flugprofil abweichende Luftfahrzeuge unter Verwendung der aktuellen Flugsicherungssysteme weiterhin koordiniert werden. Lokale Besonderheiten können durch Anpassung der Taskzeiten berücksichtigt werden.

Die genannten Tasks werden in eine Formel umgesetzt, um die Workload in Minuten zu berechnen. Diese Berechnungsvorschrift ist in dem Tool zur Analyse der Kapazitätsanforderungen und dem Schnellzeitsimulator eingefügt worden.

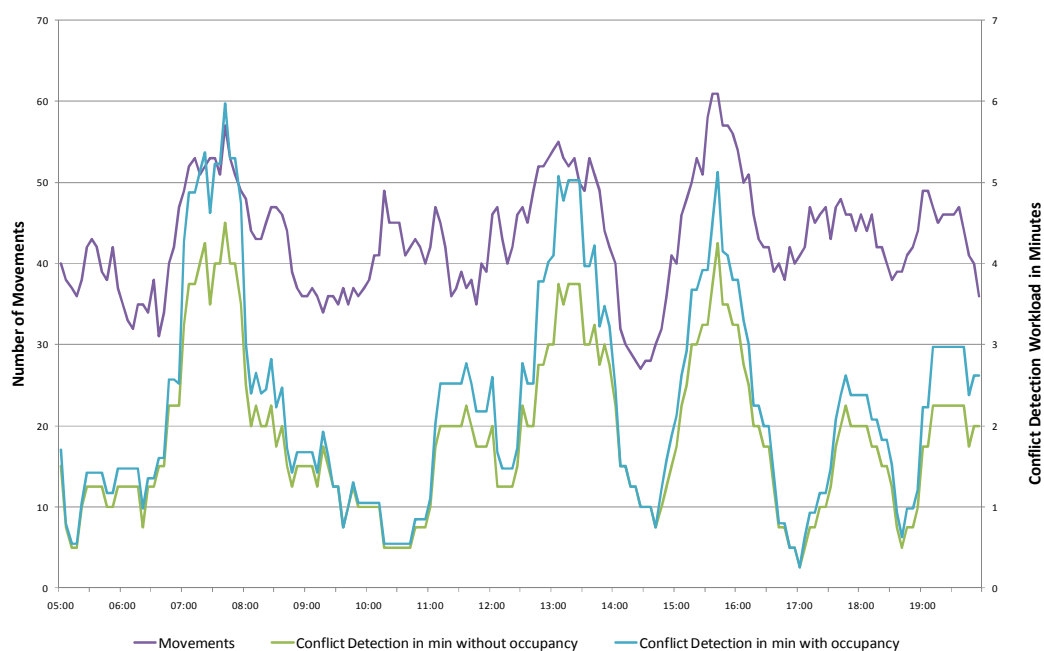


Abbildung 3: Vergleich der Arbeitslast „Konfliktsuche“ mit und ohne Occupancy

3.8 Vergleich der Workload Modelle Eventbasiert gegen Taskbasiert:

Im Laufe der Testphase im Rahmen von Schnellzeitsimulationen wurden verschiedene Verkehrsbeispiele, die reale Verkehrssituationen abbilden, mit jeweils hohen Spitzenbelastungen untersucht. Die Arbeitslastergebnisse wurden bestimmt und mit den Ergebnissen des eventbasierten DFS Workload-Modells verglichen. Im Ergebnis wurde eine hohe Übereinstimmung

zwischen der genutzten 90 Punkt-Arbeitslastgrenze und der verwendeten 70% Belastungsschwelle festgestellt. In Abbildung 4 ist ein Vergleich für einen Sektor grafisch dargestellt.

Da die Modelle (Arbeitslastpunkte gegen Zeitdauer) nicht direkt vergleichbar sind, wird als Basis das prozentuale Verhältnis zum Arbeitslastgrenzwert miteinander verglichen. Das eventbasierte Modell wurde in den letzten Jahren immer wieder mit dem System KARLDAP (streifen-

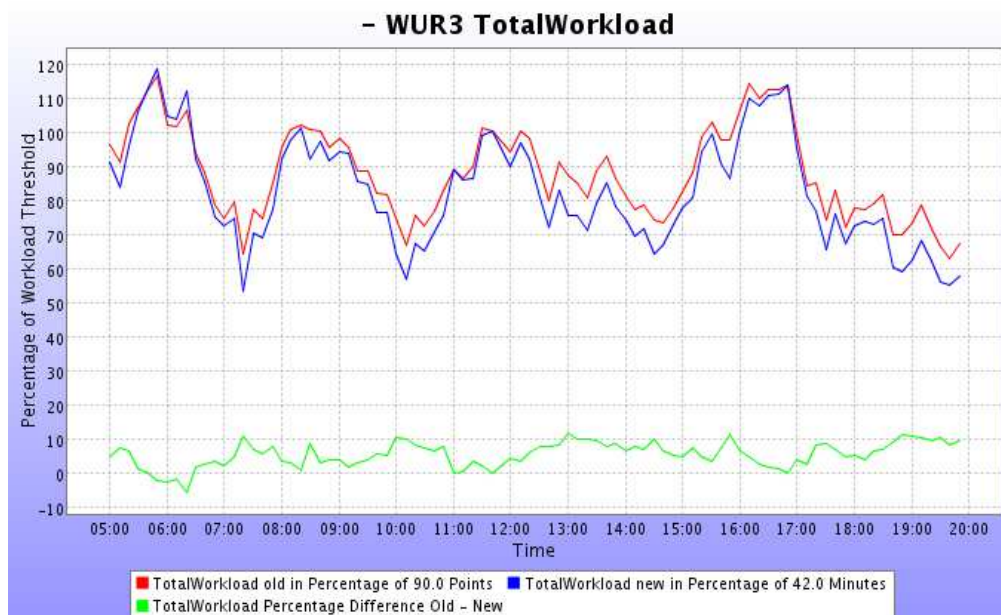


Abbildung 4: Prozentualer Vergleich der Gesamtarbeitslast zwischen eventbasiertem und taskbasiertem Arbeitslastmodell

basiertes System) validiert. Für das taskbasierte Modell wurden in diesem Vergleich Parameter aus Studien und Anwendungserfahrungen innerhalb der DFS für streifenbasierte ATC-Systeme verwendet.

3.9 Methode zur Parametererfassung

Die Zeitparameter der einzelnen Tasks wurden im ersten Schritt aufgrund von Anwendungserfahrungen definiert. Für das neue ATC-System P1/VAFORIT musste aber ein angepasster Parametersatz angewendet werden. Um die Parameter zu erhalten, fand eine Expertenrunde mit Fachleuten aus der Niederlassung Karlsruhe statt, die Zeitdauern für einzelne Tätigkeiten abgeschätzt haben. Diese Parameter wurden anschließend mit Hilfe einer Realzeitsimulation am P1/VAFORIT-System validiert. Dazu wurden die Zeitdauern einzelner Lotsentätigkeiten erfasst und in einem weiteren Schritt die zugehörigen Parameter berechnet.

Die Anforderungen und die Durchführung der Parametererfassung erforderte ein detailliertes Untersuchungskonzept, das im Folgenden dargestellt wird.

Es wurden verschiedene Sektoren untersucht. In jedem Sektor wurden jeweils die Planungs- und Radar-Position erfasst. Ergänzend zur Beobachtung erfolgten Aufzeichnungen von Simulationsdaten. Zusätzlich wurde die subjektive Einschätzung der Arbeitsbeanspruchung der Lotsen während des Runs abgefragt. Es fand eine systematische Untersuchung der einzelnen Task-Parameter anhand ausgewählter Szenarien mit verschiedenen Verkehrsmengen statt. Die Auswertung erfolgte vollständig anonymisiert, und es wurden keine personenbezogenen Daten erfasst.

Die Dauer einer einzelnen Simulation betrug 1:15 h. Bei jedem Simulationslauf wurde der Messzeitraum von 1:00 h eingehalten. In der gesamten Simulationskampagne wurden 7 Szenarien mit unterschiedlichen Verkehrsmengen eingesetzt. 5 Szenarien hatten eine mittlere bis hohe Verkehrslast (50-60 Flugzeuge pro Sektor und Stunde) und 2 Szenarien hatten eine sehr hohe Verkehrsmenge (78 Flugzeuge pro Sektor und Stunde). Es wurden 3 Sektoren aus 3 Einsatz-Berechtigungs-Gruppen (EBG) im Verbund simuliert. Es handelt sich um die Sektoren Tango (TGO), Würzburg (WUR) und Erlangen (ERL),

die sowohl in der sogenannten Low-Konfiguration als auch in der High-Konfiguration untersucht wurden. Es wurden in jeder Konfiguration 5 Simulationsläufe durchgeführt.

Als Teilnehmer standen 10 Lotsen mit vertieften Erfahrungen des P1/VAFORIT-Systems und 8 Lotsen als Beobachter bereit. Die Beobachter hatten ebenfalls operative P1/VAFORIT-Kenntnisse. Die Teilnehmer hatten Berechtigungen der jeweils beteiligten EBGs, der Luftraum war daher jedem bekannt. Die Beobachter wurden für ihre Aufgabe eigens geschult und alle Teilnehmer in die Untersuchung und das Verfahren eingewiesen. Die Anzahl und Dauer einer relevanten Tätigkeit wurde mit einem Erfassungstool aufgenommen. Durch die während der Untersuchungstage erlangte Übung bzw. Routine (Beobachter wie Lotse) konnten die Beobachtungen zunehmend gezielter und genauer erfasst werden. Der Lotse teilte den Beginn und das Ende einer Tätigkeit durch verbale Äußerung von Schlüsselworten mit. Durch diese besondere Methodik können die nicht sichtbaren mentalen Aktivitäten der Lotsen transparent gemacht werden und sie ermöglicht damit einen Einblick in seine Denkprozesse. Dies ist insbesondere bei kognitiven Aktivitäten wie Konfliktsuche und Konfliktlösung notwendig.

Nach jedem Simulationslauf wurden besondere Ereignisse und Auffälligkeiten in einem Simulationsprotokoll festgehalten (Leemüller & Paul, 2011).

Zur Erzielung von belastbaren Daten wurden mehrere Variationen erzeugt durch

- Vermeidung von Wiederholungen von bekannten Verkehrsbeispielen,
- Szenarien für jeweils Low und High Konfiguration der Sektoren für 5 Simulationsläufe,
- Wechsel der Positionen Radar und Planung,
- Verifikation durch zwei High-Traffic-Szenarien für Low und High Konfiguration der Sektoren,
- Positionswechsel der Beobachter,
- Einsatz von mehreren Teilnehmern.

3.10 Datenauswertung

Es wurden insgesamt 8333 Taskbeobachtungen an allen Radar-Positionen und 3204 Messwerte

an allen Planungs-Positionen erfasst. Zur Auswertung wurden zunächst unplausible Werte entfernt und anschließend eine statistische Ausreißer-Analyse durchgeführt. Die Auswertung der erfassten Daten wurde mit aufgezeichneten Simulationsdaten ergänzt.

Eine detailliertere Darstellung der Task-Dauern nach Sektoren, Konfiguration und Radar- oder Planungs-Lotse sind in den Abbildung 5 und Abbildung 6 gezeigt.

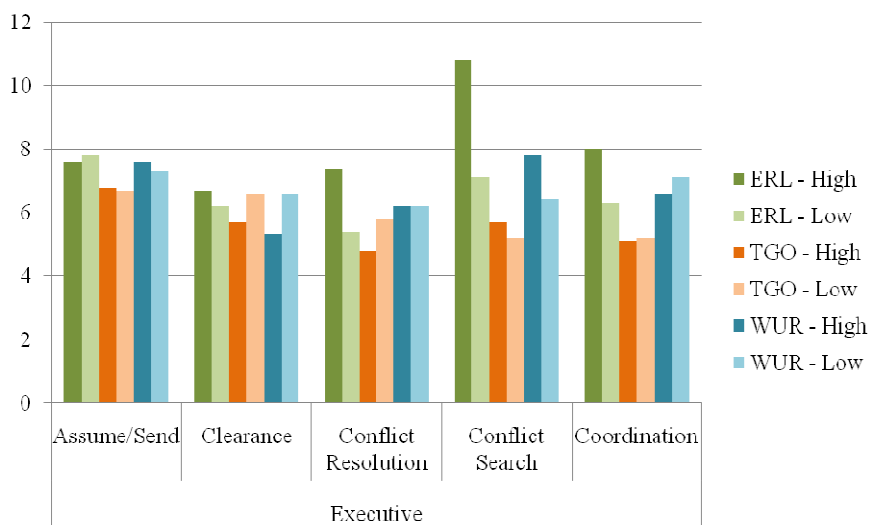


Abbildung 5: Durchschnittliche Dauer der beobachteten Aktivitäten – Radarlotse

Betrachtet man die einzelnen Task-Dauern für den Radarlotse in Abbildung 5, so stellt man fest, dass Routinetätigkeiten wie „Assume/Send“ und „Clearance“ sehr einheitlich sind; die Unterschiede zwischen den Sektoren wie auch zwischen der Konfiguration eines Sektors fallen diesbezüglich sehr gering aus. Auffällig sind hin-

gegen die Unterschiede im ERL-Sektor zwischen High und Low-Konfiguration bei den Task-Dauern „Conflict Search“ und „Conflict Resolution“ sowie „Coordination“. Die anderen Sektoren erscheinen dagegen homogener.

In Abbildung 6 sind die Task-Dauern für den Planungslotse dargestellt. Auffällig sind die Unterschiede zwischen High und Low-Konfiguration in allen Sektoren bei den Tasks „Conflict Resolution“ und „Conflict Search“. Hier fallen die Zeiten in der High-Konfiguration z.T. deutlich höher aus als in der Low-Konfiguration. Bei den Coordination Tasks „Electronic“ oder „Telephone“ ist das Bild uneinheitlich. Im Sektor ERL fallen diese für High größer aus als für Low; im TGO-Sektor ist es umgekehrt, die Low-Sektoren haben längere Taskzeiten, während es im Sektor WUR unsystematisch ist. Die Auswertung der Koordinationsdaten berücksichtigt dieses asymmetrische Ergebnis.

Die Taskzeiten für das Monitoring können nicht in einer Realzeitsimulation erfasst werden, sondern werden aus Studien entnommen, die durch frühere Anwendungserfahrungen in der DFS bestätigt wurden.

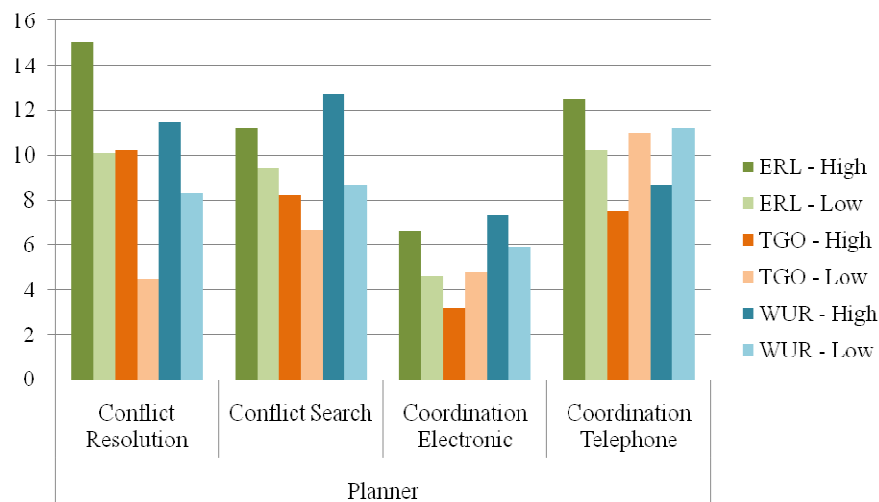


Abbildung 6: Durchschnittliche Dauer der beobachteten Aktivitäten – Planungslotse

3.11 Subjektive Beanspruchung während des Simulationslaufes

Die subjektive Beanspruchung während des Simulationslaufes wurde ebenfalls erhoben. Als Methode wird das standardisierte Verfahren „Instantaneous Self Assessment“ (ISA) verwendet. Das Verfahren beruht auf einer 5-stufigen Skala. Alle 10 Minuten ertönte ein Signal, und der Lotse sollte daraufhin eine Einschätzung seiner Beanspruchung verbal abgeben. Dieser Wert wurde auf dem Simulationsprotokoll notiert.

Aufgrund der kontinuierlichen Datenerhebung ermöglicht das Verfahren die Erstellung eines zeitlich fortschreitenden Abbildes der Arbeitsbeanspruchung des Lotsen, sodass bei einer nachträglichen Betrachtung der Ergebnisse eine direkte Zuordnung möglicher Auffälligkeiten zu bestimmten Ereignissen erfolgen kann.

Die Auswertung der subjektiven Beanspruchung ergab folgende Ergebnisse über alle Simulationsläufe aufgeschlüsselt nach Sektoren und getrennt nach Planungs- und Radar-Position (Abbildung 7). Die Beanspruchung des Planungslotsen fiel in allen Sektoren geringer aus als die des Radarlotsen aus. Es ist außerdem zu erkennen, dass die Beanspruchung in den drei Sektoren unterschiedlich bewertet wurde, was auf die jeweilige Verkehrsmenge zurückzuführen ist, die durch die gestrichelte Linie dargestellt ist.

Eine Berechnung der Korrelation nach Pearson zwischen subjektiver Beanspruchung und Verkehrsmenge in den 10 min Messintervallen ergab einen Korrelationskoeffizienten von 0,70 für die Radar-Position und 0,66 für die Planungs-Position.

4 Ergebnisse

Für das taskbasierte Workload-Modell liegen drei Parametersätze vor. Diese wurden, wie beschrieben, aus Studien, aus Expertenworkshops mit Lotsen und durch die Kalibrierung mittels Realzeitsimulationen gewonnen. Es wurde für jeden Parametersatz die Workload-Kurven im Rahmen von Schnellzeitsimulationen berechnet.

Dazu wurde ein Szenario basierend auf dem realen Verkehr eines Tages herangezogen. Die Workload-Kurven der verschiedenen Parameter sind in Abbildung 8 vergleichend für einen Sektor beispielhaft dargestellt. Die 42-Minuten-Grenze entspricht der 70% Maximalbelastung. Dass die Maximalbelastung von 60 Minuten im oberen Diagramm mehrfach überschritten wurde, liegt daran, dass in der Simulation der gesamte Sektor zugrunde gelegt wurde, während in der Realität der Sektor gesplittet wurde.

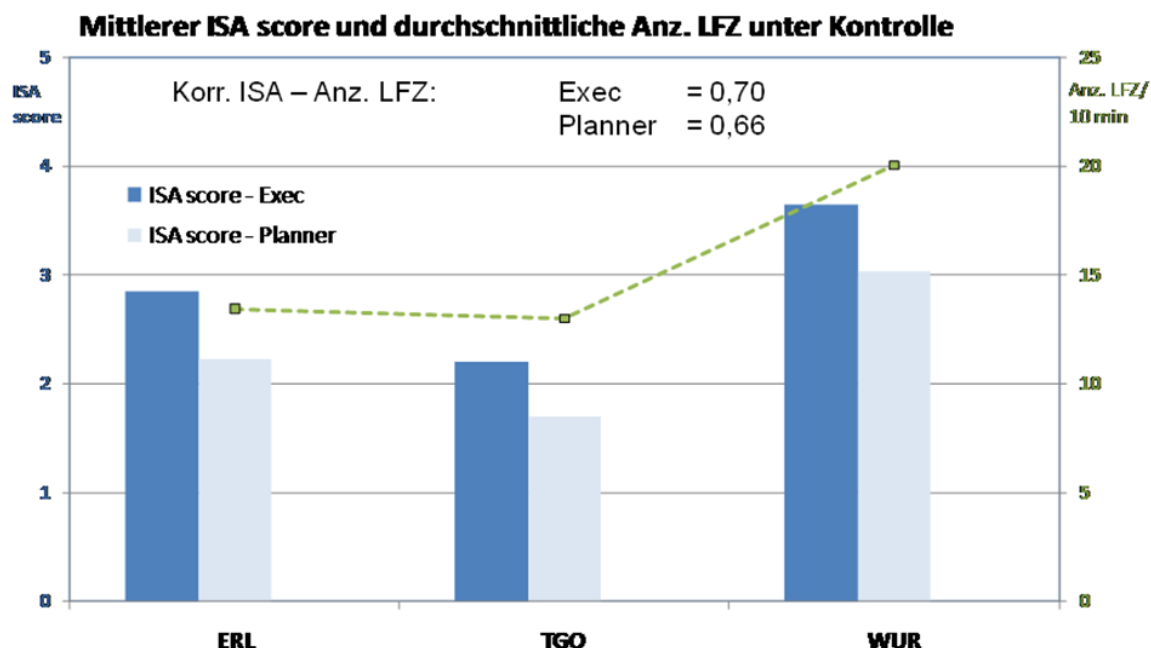


Abbildung 7: Mittlere subjektive Beanspruchung

Anhand der Kurven für den Radarlotsen (Abbildung 8, oben) erkennt man, dass die Studienparameter und Expertenmeinung ähnliche Beanspruchungswerte liefern. Die Kalibrierungsergebnisse fallen dagegen niedriger aus.

Für den Planungslotsen (Abbildung 8, unten) zeigt sich dagegen, dass die Kurve mit den Parametern aus Studien die niedrigsten Beanspruchungswerte aufweist. Die Kurven mit den Experten-

und Kalibrierungsparametern sind in dieser Situation fast gleichwertig.

Der Vergleich zwischen Radarlotse und Planungslotse ergibt somit ein uneinheitliches Bild. Die Workload-Kurve basierend auf Studienparametern zeigt beim Radarlotsen die größte Beanspruchung, beim Planungslotsen dagegen die

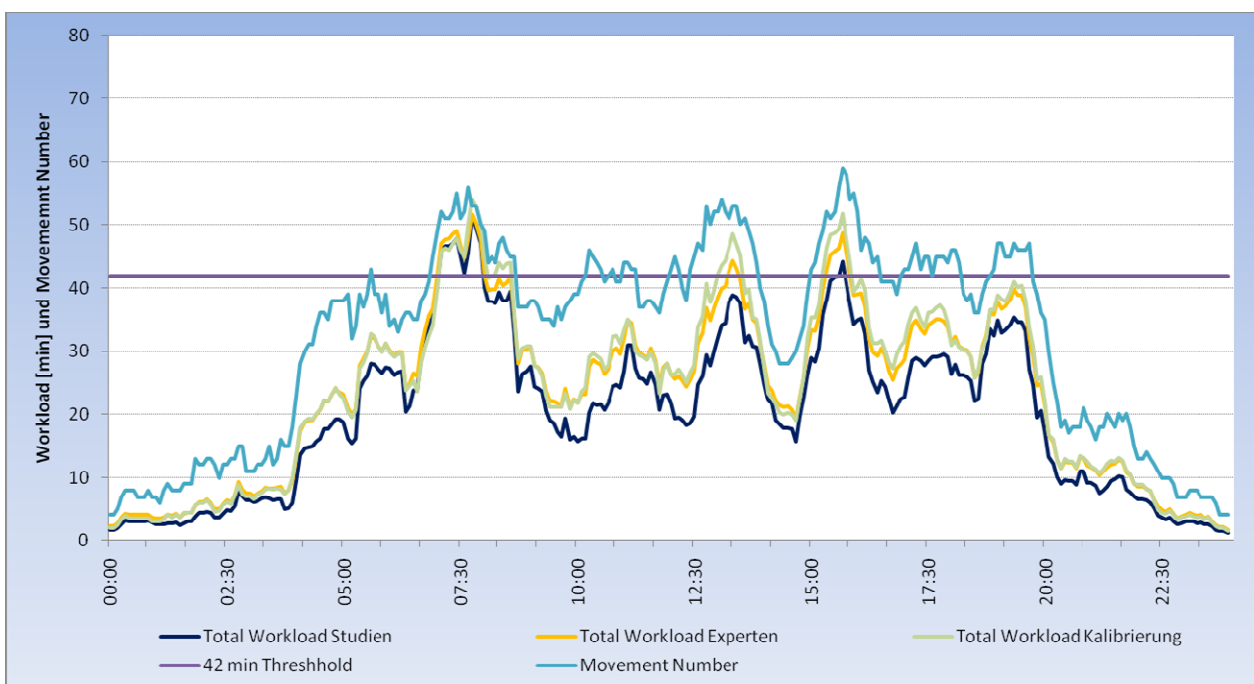
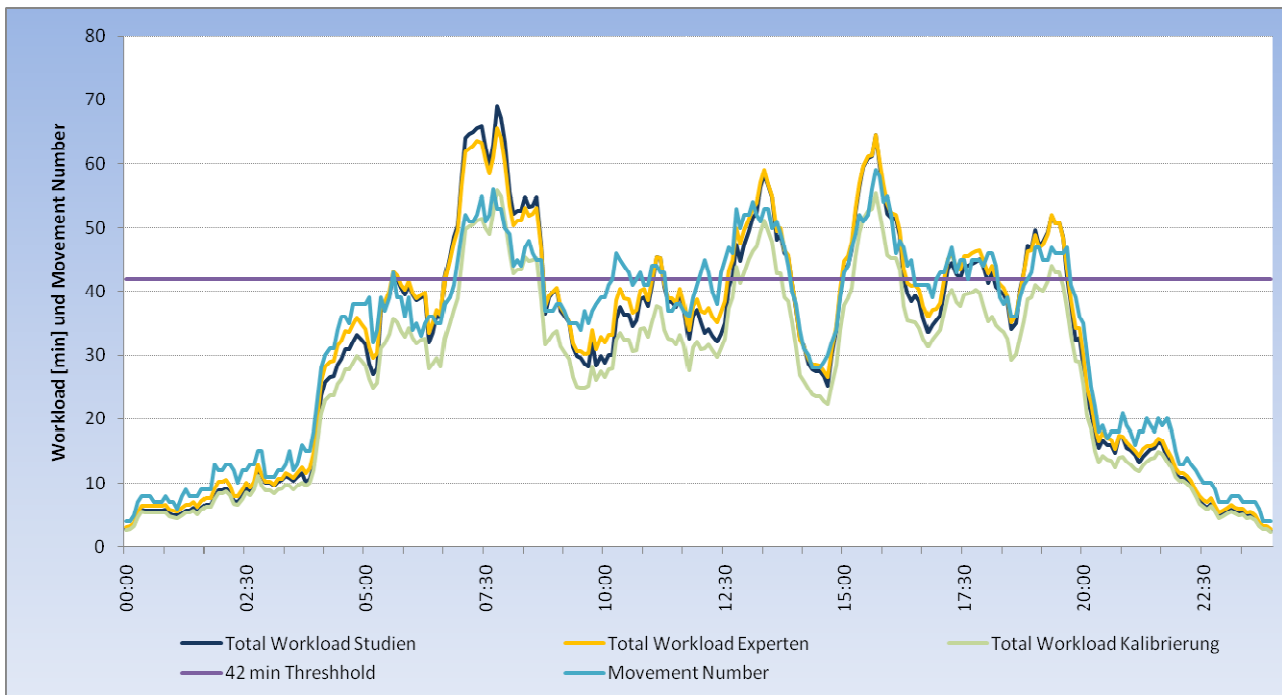


Abb. 8: Workload Kurven-Vergleich Sektor FFMB; oben Radarlotse, unten Planungslotse

geringste. Die Workload-Kurve aufgrund der kalibrierten Parameter ist beim Radarlotsen am niedrigsten und beim Planungslotsen am größten. Dieses Ergebnis ist aber verständlich, wenn man das zu Grunde gelegte ATC-System mit einbezieht. Die Parameter aus den Studien sind bei streifenbasierten Systemen erhoben worden. Die Kalibrierung hingegen wurde bei einem trajektorien-basierten System durchgeführt. Bei diesen Systemen wird die Beanspruchung innerhalb des Lotsenteams hin zur Planung verschoben.

Die Abschätzung der Beanspruchung durch Experten ist beim Planungslotsen fast gleich wie die Kalibrierungsergebnisse. Daran erkennt man, dass die Expertenmeinungen eigentlich zuverlässige Vorhersagen liefern. Beim Radarlotsen allerdings zeigt die Expertendefinition eine viel höhere Beanspruchung als die Kalibrierung. Die Lotsen haben hier eine konservativere Einschätzung vorgenommen, um kein Risiko bei der Sicherheit einzugehen. Der Radarlotse ist in dem Team immer noch die Position mit der höheren Beanspruchung.

Es zeigt sich also, dass bei Einführung eines neuen ATC-Systems mit einer neuen Bedienphilosophie immer eine Kalibrierung der Taskparameter notwendig ist. Die Verschiebung der Arbeitsbelastung zum Planungslotsen lässt einen Kapazitätsgewinn erwarten, der durch die Kalibrierungsuntersuchung erkennbar wurde. Das bisherige eventbasierte Workload-Modell konnte das trajektorien-orientierte ATC-System nicht adäquat abbilden.

5 Vergleich der subjektiven Beanspruchung mit dem Workloadmodell

Wie in Abschnitt 2 erläutert, gibt es verschiedene Möglichkeiten die Arbeitsbeanspruchung von Fluglotsen zu bestimmen. Zum einen durch direkte Messungen und zum anderen indirekt durch Berechnungen mit geeigneten Workloadmodellen. Wie ein solches Workloadmodell entwickelt und kalibriert wird, wurde in den vorherigen Abschnitten erläutert. Nun gilt es abschließend zu klären, wie gut das Workloadmodell mit der subjektiven Beanspruchung von Lotsen übereinstimmt. Man spricht hierbei auch von einer Validierung, d.h. dem Nachweis, dass eine analytische Methode für ihren Einsatzzweck geeignet ist.

Im Rahmen einer Bachelorarbeit (Emsbach, 2011) wurde der Zusammenhang der subjektiven Beanspruchung mit dem Workload aus dem CAPANneu Modell untersucht.

Zur Gegenüberstellung der erhobenen ISA-Daten mit den in Schnellzeitsimulationen zu ermittelnden Workloadwerten gemäß der Workloadformel, musste vorab eine Auswahl der zu vergleichenden Runs erfolgen, um repräsentative Aussagen über deren Korrelation treffen zu können. Da einige Verkehrsszenarien in der Realzeitsimulation mehrfach verwendet wurden, wurden diese für einen Vergleich herangezogen. Es wurden drei Szenarien, zwei Low und eine High Konfiguration, aus der Echtzeitsimulation verwendet, sodass Simulationsdaten von insgesamt sechs Runs verfügbar waren. Damit ist stichprobenartig ein Vergleich der Beanspruchungswerte aus Realzeit- und Schnellzeitsimulation möglich. Auf diese Weise können die Workloadkurven einer Schnellzeitsimulation mit den ISA-Werten aus jeweils zwei Echtzeitsimulationen verglichen werden.

Die Realzeitszenarien mussten allerdings an die Erfordernisse der Schnellzeitsimulation angepasst werden, d.h. Flugplandaten in ein Format umgewandelt werden, dass in den Schnellzeitsimulator AirTop geladen werden kann. Ferner mussten noch die Startzeiten der Flüge optimiert werden, damit vergleichbare Verkehrszahlen in den zehn-Minuten Intervallen zwischen Realzeit- und Schnellzeitsimulation vorlagen.

Die aus dieser Simulation gewonnenen Daten werden daraufhin in den Java Reporter gegeben, der daraufhin, gemäß der Workloadformel, die Arbeitslast der Lotsen errechnet.

Zum Vergleich der ISA-Werte mit den Workloadberechnungen mussten die verfügbaren Messintervalle aus der Realzeitsimulationen mit denen der Schnellzeitsimulation übereinstimmen. Während eine Schnellzeitsimulation, die mit einem bestimmten Verkehrsszenario durchgeführt wird, immer das gleiche Ergebnis erzeugen wird, gilt das für Realzeitsimulationen nicht. Wird dasselbe Verkehrsszenario zweimal durchgeführt, kann es jeweils unterschiedliche Ergebnisse liefern, je nachdem wie der Lotse gearbeitet hat, d.h. den Verkehr gesteuert, potentielle Konflikte gelöst und Interaktionen mit Piloten und Lotsen anderer Sektoren durchgeführt hat.

Eine Methode um Daten miteinander vergleichen zu können und den linearen Zusammenhang festzustellen, bietet das Erstellen von Korrelationen. Dieses ist ein Verfahren, welches einen linearen Zusammenhang zwischen zwei Datenmengen anhand von einer dimensionslosen Skalierung zwischen -1 und +1, erstellt. Dabei bildet ein Korrelationskoeffizient von 0 keinen linearen Zusammenhang ab, während bei einem Wert von +1 (bzw. von -1) ein vollständig positiver (bzw. negativer) Zusammenhang besteht.

Um sicherzustellen, ob die gewählten Intervalle innerhalb der jeweiligen Runs in der Schnellzeitsimulation mit denen der Echtzeitsimulation weitestgehend übereinstimmen, wurden Korrelationen nach Pearson der Flugverkehrszahlen der beiden Simulationen miteinander gerechnet. Die Berechnung zwischen den beiden Flugzahlen in den zehn- minütigen Messintervallen ergab einen Korrelationskoeffizienten von $r = 0,87$. Aus diesem Grund kann von einem sehr hohen Zusammenhang zwischen den Flugbewegungen aus Schnellzeit- und Echtzeitsimulation ausgegangen werden. Dies deutet auf eine geeignete Wahl der Intervalle und auf eine gute Darstellung der realen Begebenheiten innerhalb der Schnellzeitsimulation hin.

Ein Vergleich zwischen den aus der Schnellzeitsimulation errechneten Workloaddaten und den aus der Echtzeitsimulation erhaltenen ISA-Scores ist demnach valide möglich.

Der ISA- Score ist ordinalskaliert, sodass sich hier die Bildung des Korrelationskoeffizienten ρ (rho) nach Spearman anbietet. Im Folgenden werden die Korrelationen, in welchen der ISA-Score beteiligt ist, mit dem Korrelationskoeffizienten nach Spearman berechnet, während alle anderen Korrelationen weiterhin auf Berechnungen mit dem Korrelationskoeffizienten nach Pearson beruhen.

In Tabelle 1 sind alle Korrelationskoeffizienten der verschiedenen Positionen und Konfigurationen dargestellt. Verglichen wurden dabei die ISA- Scores mit der errechneten Workload und der Anzahl an Luftfahrzeugen, sowie die Anzahl Luftfahrzeuge aus der Schnellzeit- mit denen der Echtzeitsimulation. Die mit einem Stern gekennzeichneten Koeffizienten sind signifikant (Irrtumswahrscheinlichkeit $p \leq 0,05$).

Korrelation zwischen:		Position	High + Low	Low	High
ISA	Capan WL	EC	0,565*	0,647*	0,363*
		PC	0,477*	0,534*	0,360*
Anzahl Lfz. RTS	ISA	EC	0,626*	0,778*	0,101
		PC	0,612*	0,741*	0,167
Anzahl Lfz. FTS	Capan WL	EC	0,913*	0,938*	0,715*
		PC	0,857*	0,903*	0,563*
Anzahl Lfz. FTS	Anzahl Lfz. RTS		0,859*	0,870*	0,714*

Tabelle 1: Übersicht aller Korrelationskoeffizienten

Vergleicht man nun die subjektive Arbeitsbeanspruchung der Lotsen aus der Low- Konfiguration in der Echtzeitsimulation mit den in der Schnellzeitsimulation errechneten Workloadwerten des Arbeitslastmodells CAPANneu, ergeben sich Korrelationskoeffizienten von $\rho = 0,65$ für den Radarlotsen und $\rho = 0,53$ für den Planungslotsen. Da Korrelationskoeffizienten größer 0,5 einen guten linearen Zusammenhang zwischen den beteiligten Größen darstellen, steht die Arbeitsbeanspruchung des Arbeitslastmodells in Zusammenhang mit der subjektiven Beanspruchung der Lotsen.

Betrachtet man nun in der High- Konfiguration den Zusammenhang zwischen der subjektiven und der errechneten Workload ergeben sich allerdings sehr geringe Korrelationskoeffizienten. Diese liegen sowohl für den Radar- als auch für den Planungslotsen bei $\rho = 0,36$ und sind somit erheblich kleiner als die Koeffizienten der Low-Konfiguration. Da der Zusammenhang zwischen subjektiver Beanspruchung und errechneter Workload in der High- Konfiguration nicht eindeutig ist, erfolgt zunächst eine Betrachtung des Zusammenhangs zwischen dem ISA- Score und der Anzahl der Luftfahrzeuge. Auf diese Weise

kann möglicherweise eine Begründung für diese schwache Beziehung gefunden werden. Stellt man also die subjektive Beanspruchung der Anzahl Luftfahrzeuge aus der Echtzeitsimulation gegenüber, erhält man ebenfalls niedrige Korrelationskoeffizienten von $\rho = 0,10$ für die Radarposition und $\rho = 0,17$ für die Planungsposition.

Eine grafische Darstellung dieser Korrelation ist in Abbildung 9 zu sehen. Es wird in dieser Abbildung deutlich, dass der Radarlotse hier zwar ebenfalls eine höhere subjektive Beanspruchung als der Planungslotse erfährt, es ist jedoch kein Zusammenhang zwischen der empfundenen Beanspruchung und der Verkehrsmenge ersichtlich, da die Verkehrsmenge über alle ISA- Werte annähernd konstant bleibt.

Aufgrund der widersprüchlichen Aussagen bezüglich der Korrelationen von der High- und Low-Konfiguration ist eine nähere Betrachtung der Unterschiede zwischen den beiden Konfigurationen angebracht. Vergleicht man die Verkehrszahlen und die errechnete Workload der beiden Konfigurationen, wie in Abbildung 10 dargestellt, wird deutlich, dass die Low- Konfiguration zum einen höhere Verkehrszahlen aufweist und zum anderen auch eine viel höhere Beanspruchung erfährt als die High- Konfiguration.

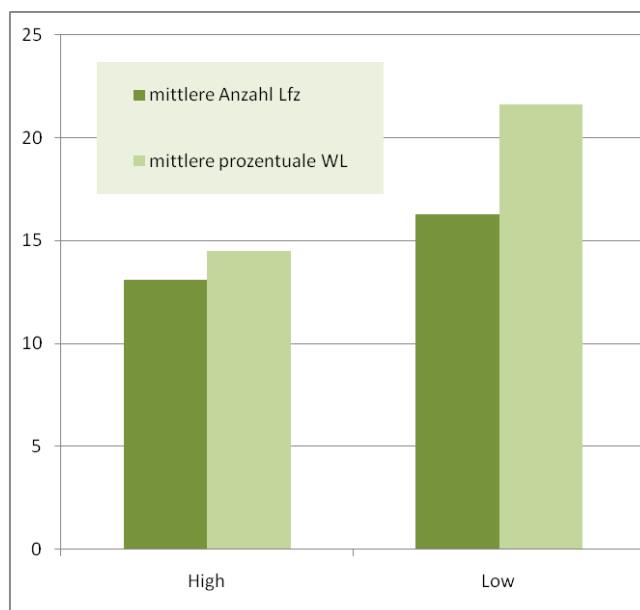


Abbildung 10: Vergleich der High- und Low- Konfiguration

Die Unterschiede der Beanspruchung zwischen den Konfigurationen sind vor allem durch die unterschiedliche Verteilung der durchzuführenden Tätigkeiten begründbar. Denn im unteren Luftraum kommt es häufiger zu Koordinationstätigkeiten bzw. zum Erteilen von Freigaben, da in diesem Luftraum häufiger vertikale Änderungen, z.B. aufgrund von An- und Abflügen zu bzw. von Flughäfen, vollzogen werden. Zudem steigt mit erhöhtem Verkehrsaufkommen die Anzahl der Tätigkeiten der Konfliktbearbeitung, welche

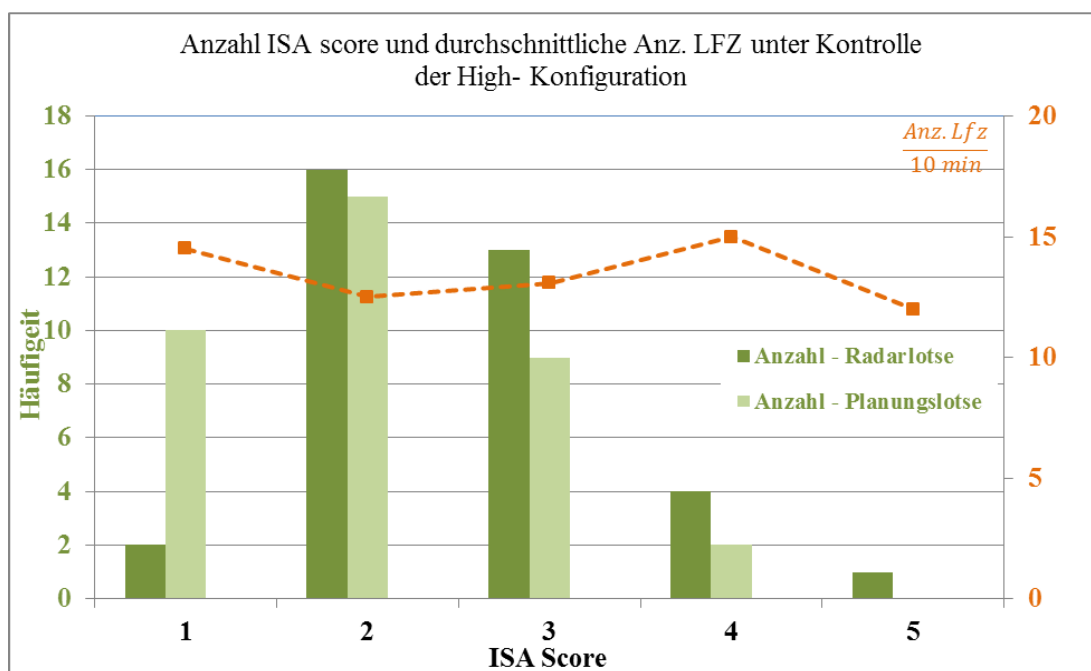


Abbildung 9: Vergleich Anzahl ISA- Score und Lfz. der High- Konfiguration

ebenfalls Einfluss auf die Arbeitsbeanspruchung der Lotsen nimmt. Die Anzahl der zu koordinierenden Luftfahrzeuge ist daher einer der ausschlaggebendsten Faktoren der Arbeitsbeanspruchung. Es nehmen allerdings auch andere, zum Teil auch äußere, Faktoren Einfluss auf die Beanspruchung der Fluglotsen. Diese sind bei einer großen Anzahl an Luftfahrzeugen jedoch eher als gering einzuschätzen, denn je größer die zu koordinierende Verkehrsmenge, desto geringer ist der Anteil dieser Störfaktoren an der gesamten Beanspruchung. Fällt die Verkehrsmenge jedoch eher gering aus, nehmen die, in der Simulation nicht erfassten, Störvariablen größeren Einfluss auf die Gesamtbeanspruchung. Zudem auch ein geringes Verkehrsaufkommen komplexe und dynamische Situationen hervorrufen kann. Da diese Störvariablen durch die Erfassung der subjektiven Beanspruchung mit berücksichtigt werden, kann es dadurch zu einer erhöhten subjektiven Beanspruchung, und somit zu höheren ISA-Scores, auch bei geringeren Verkehrsmengen kommen. Ein zusätzlicher Grund könnte auch eine Fehleinschätzung des ISA-Scores durch die Lotsen bei der High-Konfiguration sein, da die Abfrage bei der High-Konfiguration als erstes eingesetzt wurde und die Lotsen noch ungeübt waren. Dies würde auch die niedrigen Korrelationskoeffizienten des Vergleiches von subjektiver und errechneter Workload begründen. Eine abschließende Aussage, warum eine Korrelation nicht nachweisbar ist, kann nicht getroffen werden. Weitere Untersuchungen und Auswertungen müssten durchgeführt werden, um die aufgetretene Situation zu klären.

Lässt man nun die Konfigurationen außer Acht und betrachtet den gesamten Luftraum, zeigt sich, trotz der niedrigen Korrelationskoeffizienten der High-Konfiguration, ein Zusammenhang zwischen der subjektiven und der berechneten Workload. Für den Radarlotsen ergibt sich ein Korrelationskoeffizient von $\rho = 0,57$ und für den Planungslotsen von $\rho = 0,48$, die beide noch als gut interpretiert werden können.

Aufgrund des Aufwandes konnten in dieser Studie nur stichprobenartige Vergleiche zwischen Echtzeitszenarien und der erhobenen subjektiven Beanspruchung mit den Ergebnissen des Workloadmodells aufgrund von Schnellzeitsimulationen durchgeführt werden. Die gefundenen Korrelationen lassen aber darauf schließen, dass der Zusammenhang zwischen subjektiver Bean-

spruchung und Arbeitslastmodell hinreichend gut ist. Es ist außerdem zu erwarten, dass in umfangreicheren Untersuchungen und damit verfügbaren Datenmaterial, diese Beziehung noch deutlicher hervortreten wird.

6 Resümée

Das entwickelte Workload-Modell dient als Grundlage für die Berechnung von Kapazitätswerten für die Sektoren der Niederlassung Karlsruhe mit dem System CAPAN. Die im Modell verwendeten Zeitparameter für die Controller-Task wurden für die P1/VAFORIT-Umgebung im Rahmen umfangreicher Realzeitsimulationen ermittelt und in das Modell integriert. Die im Anschluss durchgeführte Schnellzeitsimulation erbrachte auf der Basis der neuen Task-Parameter plausible Ergebnisse für die Kriterien Belastung und Kapazität. Das erarbeitete Modell einschließlich der ermittelten Parameter wurde in die neue CAPAN-Software implementiert und steht für die Ermittlung von Belastungen und Sektorkapazitätswerten zur Verfügung.

Zudem konnte in Stichproben eine gute Übereinstimmung der in Schnellzeitsimulationen berechneten CAPAN Workloadwerten und der subjektiven Beanspruchung aus den Realzeitsimulationen festgestellt werden, die zeigen, dass das Workloadmodell valide ist.

Es ist geplant, die Zeitparameter nach einer Eingewöhnungszeit der Lotsen an das VAFORIT-System zu überprüfen und mit den Berechnungen abzugleichen. Aufgrund von Änderungen in der Verkehrszusammensetzung, Verfahrensanpassungen und Verbesserungen des ATC-Systems ist eine periodische Überprüfung der Taskparameter empfehlenswert. Das Workload-Modell wird auch für den unteren Luftraum und Approach kalibriert. Somit steht demnächst ein einheitliches Modell für alle Niederlassungen der DFS zur Verfügung, das sich durch eine hohe Transparenz auszeichnet und den Anforderungen des Betriebs gerecht wird.

Literatur

Emsbach, T (2011). *Taskbasiertes Arbeitslastmodell für Flugsicherungssysteme* (Bachelorarbeit). Aachen: Fachhochschule, Luft- und Raumfahrttechnik.

Leemüller, R. & Paul, B. (2011). *Entwicklung des taskbasierten Arbeitslastmodells für die Nutzung in CAPAN der Niederlassung Karlsruhe* (unveröff. Bericht). Langen: DFS Deutsche Flugsicherung GmbH.

Salas, M. (2010). *Human Factors in Aviation*. New York: Elsevier Inc.

Udovic, A. & Tenoort, S. (2001). *Methoden zur Berechnung von Kapazität und Workload* (unveröff. Bericht). Langen: DFS Deutsche Flugsicherung GmbH.

Impressum:

TE im Fokus - Informationen aus dem Bereich Forschung und Entwicklung der DFS Deutsche Flugsicherung GmbH

TE im Fokus erscheint in der Regel halbjährlich. Dieser Bericht ist elektronisch im Internet (www.dfs.de, Modul Forschung & Entwicklung → Forschungszeitschrift) sowie über das DFS Intranet, Dokumentationsbereich im F&E Portal (<http://forschungszentrum.lgn.dfs.local>) verfügbar. 70 Exemplare werden in gedruckter Form aufgelegt.

Herausgeber: Bereich Forschung und Entwicklung, TE
DFS Deutsche Flugsicherung GmbH
Am DFS-Campus 5
63225 Langen

Redaktion: Dr. Andreas Herber (andreas.herber@dfs.de),
Oliver Haßa (oliver.hassa@dfs.de),
Stefan Tenoort (stefan.tenoort@dfs.de)

Sekretariat: Petra Schuster (petra.schuster@dfs.de)
DFS Deutsche Flugsicherung
Telefon: 06103-707-5751
Telefax: 06103-707-5741

Dieser Zeitschrift ist eine ISSN (International Standard Serial Number) zugeteilt. Dabei ist zu unterscheiden zwischen der Print- und der Internet-Version. Diese sind wie folgt:

Printversion: ISSN 1861-6364

Internet-Version: ISSN 1861-6372

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 © 2011 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.