

Entwicklung und Untersuchung einer Stresstestmethode für die Risikoanalyse von soziotechnischen Systemen in der Luftfahrt

Dissertation

zur Erlangung des akademischen Grades

Doktoringenieur

(Dr. Ing.)

vorgelegt

der Fakultät Verkehrswissenschaften „Friedrich List“

der Technischen Universität Dresden

von

Dipl.-Ing. Lothar Meyer

Geboren am 24. Oktober 1981 in Hannover

Dresden 2017

Gutachter:

Prof. Dr.-Ing. habil. Hartmut Fricke

Prof. Dr. rer. nat. habil. Sebastian Pannasch

Eingereicht am 3. Januar 2017

Verteidigung am 11. Juli 2017

Inhaltsverzeichnis

1	Wissenschaftliche Zielsetzung und Lösungsansatz	1
1.1	Einleitung	1
1.2	Der menschliche Fehler in der Luftfahrt	7
1.2.1	Der statistische Unfallbeitrag	7
1.2.2	Modelle des menschlichen Unfallbeitrags	10
1.3	Die Probleme gegenwärtiger Methoden der Risikoanalyse	11
1.3.1	Die „Safety Assessment Methodology“	13
1.3.2	Modellbasierte Methoden	14
1.3.3	Die experimentelle Stresstestanalyse	16
1.3.4	Die Unfalluntersuchung	16
1.3.5	Das „Accident-Incident-Model“	21
1.3.6	Die Problemanalyse	21
1.3.7	Die Schlussfolgerung	23
1.4	Die Echtzeitsimulation als mögliche Lösung	24
1.4.1	Der gegenwärtige Nutzen von Echtzeitsimulatoren im Flugbetrieb	24
1.4.2	Ein Konzept zur simulationsgestützten Risikoanalyse	26
1.4.3	Mögliche Einschränkungen bei der Nutzung der Echtzeitsimulation	33
1.4.4	Anwendungsszenarien der simulationsgestützten Risikoanalyse	37
1.5	Die Zielsetzung und Struktur der Arbeit	39
1.5.1	Die Problemdefinition und der Lösungsansatz	39
1.5.2	Die Untersuchungshypothese	41
1.5.3	Die Struktur der Arbeit	42
2	Konzeptentwicklung „Accelerated Risk Analysis“	43
2.1	Die grundlegenden Begriffe	44
2.1.1	Die Wahl der Risikometrik	44
2.1.2	Die Wahl der Zielsicherheit	46
2.1.3	Die Wahl der Grundgesamtheit	46
2.1.4	Die Wahl der Unfallereignisverteilung	47
2.2	Das Sicherheitsargument	47
2.3	Die Abhängigkeit der Irrtumswahrscheinlichkeit zur Anzahl der Stichproben	48
2.3.1	Die Wahl der Irrtumswahrscheinlichkeit	48
2.3.2	Die Schätzung der Abhängigkeit anhand des Tschebyscheff-Ansatzes	48
2.3.3	Die Schätzung der Abhängigkeit anhand des Binomialansatzes	49
2.4	Ein Konzept zur Beschleunigung der Konvergenz	52
2.4.1	Die Beschleunigung durch Stresseinwirkung	52
2.4.2	Die Beschleunigung nach Tschebyscheff	53
2.4.3	Die Beschleunigung nach Binomialverteilung	53
2.5	Ein Konzept zur Stimulation sicherheitsrelevanter Ereignisse	54
2.5.1	Das Ziel der Stimulation	54
2.5.2	Die Wahl des Zeitdrucks als Stressor	54
2.5.3	Der Status Quo der Verfahren zur Induktion von Zeitdruck	57
2.5.4	Ein Verfahren zur Induktion von Stress	60

2.6	Die Regressionsanalyse zur Verifikation der Zielsicherheit.....	64
2.6.1	Die Regressionsanalyse zur Bestimmung der Unfallereignisverteilung	64
2.6.2	Die Regressionsanalyse zur Verifikation der Zielsicherheit	66
2.7	Die Verifikationsziele	71
3	Experimentelle Studie	73
3.1	Das Konzept der Studie.....	75
3.1.1	Die Wahl der unabhängigen und abhängigen Größen.....	75
3.1.2	Die Kausalhypothesen	76
3.1.3	Die Anforderungen an die Pilotstudie	76
3.1.4	Die Wahl der statistischen Tests	78
3.2	Das experimentelle Design.....	79
3.2.1	Die Auswahl der Arbeitsposition	79
3.2.2	Die Auswahl der Risikometrik.....	81
3.2.3	Die Definition der Primäraufgabe	82
3.2.4	Der Surface Movement Manager	85
3.2.5	Die Implementierung des Verfahrens „Konkurrenzdruck“	86
3.2.6	Das Messverfahren	99
3.2.7	Die Befragung	101
3.2.8	Die Szenarientwicklung	102
3.3	Die Steuerung des Wettbewerbs.....	106
3.3.1	Die Reaktion der Aktivzeit	107
3.3.2	Die Reaktion der Anzahl aktiver Verkehrsbewegungen	107
3.3.3	Das Verifikationsergebnis	108
3.4	Die Organisation der Durchführung	109
3.4.1	Die Akquisition der Probanden.....	109
3.4.2	Die Versuchsplanung.....	110
3.5	Die Kalibrierung der Basislast	110
3.6	Die Kalibrierung der Zeitdrucklastszenarien	113
3.6.1	Die Analyse der Reaktionszeit.....	114
3.6.2	Die Analyse des Zeitfehlers.....	116
3.6.3	Die Analyse der subjektiven Arbeitsbeanspruchung und des Zeitdrucks	117
3.6.4	Die Analyse der Runway Incursion.....	118
3.6.5	Die Bestimmung der Zeitdrucklastparameter	119
3.7	Die Analyse der Lastenabhängigkeit der Aktivzeiten	119
3.8	Die Analyse der Lern- und Müdigkeitseffekte.....	121
3.8.1	Die Lerneffekte	121
3.8.2	Die Erschöpfungseffekte	121
3.9	Die Analyse der Stressreaktionen.....	122
3.9.1	Die Reaktionszeit	122
3.9.2	Die Aktivzeit.....	122
3.9.3	Das angepasste Zeitbudget.....	123
3.9.4	Die Arbeitsbeanspruchung und der subjektive Zeitdruck	123
3.9.5	Die Runway Incursion und der Zeitfehler	124
3.10	Diskussion der Ergebnisse	129

4	Abschlussdiskussion	133
4.1	Das Fazit über das Konzept „Accelerated Risk Analysis“	133
4.1.1	Die Effektivität des Verfahrens und mögliche Anwendungsszenarien	134
4.1.2	Rückschluss auf die Modelle zur Bestimmung des Risikos	138
4.1.3	Implikationen durch Novizen- und Expertenprobanden	139
4.2	Die Anwendung auf andere Arbeitspositionen des Flugbetriebs	140
4.3	Ein Erklärungsansatz mit dem Contextual Control Modell	142
4.4	Ein Modell zur Beschreibung von Risikobereitschaft und Schnelligkeit	143
4.5	Mögliche Hypothesen für eine Folgestudie	144
4.6	Schlusswort	145
	Abkürzungsverzeichnis	147
	Formelzeichenverzeichnis	149
	Abbildungsverzeichnis	153
	Tabellenverzeichnis	155
	Literaturverzeichnis	157
	Danksagung	163
	Anhang	165

Kurzfassung

Die Unfallstatistik der kommerziellen Luftfahrt verzeichnet in diesem Jahrzehnt einen Höchststand betrieblicher Sicherheit. Es ist das Ergebnis einer jahrzehntewährenden Entwicklung der Luftfahrt, Vor- und Unfälle systematisch zu dokumentieren und aus ihnen lernen zu können, um zukünftig Unfälle zu vermeiden. Zwangsläufig führt diese Sicherheitsaffinität zu einer Überregulierung, die den Flugbetrieb in seinem bestehenden Konzept zunehmend konserviert und die Einführung von Innovation zunehmend hemmt. Einer der Gründe für diesen Trend ist, innovative Konzepte und Systeme nicht gefahrenlos während der Konzept-, Implementierungs- und Migrationsphase des Lebenszyklus hinsichtlich des Risikos analysieren zu können. Die Erlangung von Sicherheitsnachweisen gestaltet sich schwierig, wenn der bestehende Flugbetrieb kein innovationsinduziertes Risiko akzeptieren kann. Insbesondere bei sogenannten soziotechnischen Systemen, die den Menschen und die Interaktionsvorgänge mit der Technik berücksichtigen, können innovative Konzepte zu neuartigen Gefahren führen, die aus einem komplexen Zusammenspiel, z.B. zwischen den Piloten und den Fluglotsen mit Verfahren sowie Technik, resultieren. Ein Nachweis unter Berücksichtigung einer Zielsicherheit, wie z.B. ein Unfall auf eine Milliarde Flugstunden vor Inbetriebnahme, ist praktisch nicht leistbar.

Diese Arbeit befasst sich mit der simulationsgestützten Risikoanalyse und liefert ein Konzept zur Nutzung von Echtzeitsimulatoren für die Erbringung des Sicherheitsnachweises. Das Konzept adressiert dabei ein statistisches Problem, seltene und unbekannte Gefahrensituationen mit Hilfe von Echtzeitsimulationen nicht oder unzureichend häufig beobachten zu können, um einen Rückschluss auf die Risiken eines neuartigen Systems ziehen zu können. Dieses Problem ist bekannt als statistische Rechtszensur. Das Kernelement des Konzeptes ist die Stressinduktion mit Hilfe von Zeitdruck, die ein hohes Maß an Kontrollierbarkeit und Reproduzierbarkeit des induzierten Stresses verspricht. Nach Vorbild eines beschleunigten Ermüdungstests sorgen die stressintensivierten Simulationsbedingungen für ein zunehmendes Fehlverhalten des Operators. Gefahrenereignisse und Vorfälle sollen aufgrund der Stressreaktion verstärkt zu beobachten sein, ohne diese direkt hervorzurufen. Nicht bekannte oder seltene Gefahren können somit beobachten werden, die ansonsten aufgrund der kleinen Eintrittswahrscheinlichkeit innerhalb der Simulationszeit nicht zu beobachten wären. Die Bestimmung des Risikos und der Sicherheitsnachweis werden unter Einbezug seltener Gefahren möglich. Bei gezieltem Anfahren von zwei bis drei Zeitdrucklasten lassen die aufgezeichneten Ereignisdaten einen Rückschluss von den beschleunigten auf die unbelasteten Bedingungen (Designstress) zu.

Eine experimentelle Studie testet eine prototypische Implementierung eines Zeitdruckinduktionsverfahrens hinsichtlich der Kontrollierbarkeit der Stressreaktion. Dies soll sowohl den Zeitdruck als geeigneten Stressor bestätigen als auch Hinweise auf die Effektivität des Induktionsverfahrens liefern. Der Versuchsaufbau umfasst die Arbeitsposition des Platzverkehrsloten am Flughafen Frankfurt am Main mit drei zu kontrollierenden Pisten, Rollwegen und Vorfeld an einem Surface Movement Manager. Es wurden drei studentische Probanden aus 14 ausgewählt, trainiert und mit Hilfe des Induktionsverfahrens getestet. Die aufgezeichneten Stressreaktionen zeigen kontrollierbare Reaktionen in der beobachteten Häufigkeit des Vorfalltyps Runway Incursion und Zeitfehler. Individuelle Stressreaktionen,

wie z.B. die individuelle Basisleistung und die Freiheitsgrade des Probanden, zwischen Risiko und Schnelligkeit abzuwägen, tragen zur Streuung der Stressreaktion erheblich bei. Ein modellbasierter Ansatz konnte eine Erklärung zu diesen Varianzen liefern und eine systematische Abhängigkeit der Stressreaktion der Probanden zum induzierten Zeitdruck verifizieren. Die Ergebnisse zeigen, dass ein lastenunabhängiges Verhältnis zwischen Risiko und Schnelligkeit für den Fall zu erwarten ist, dass der Operateur höhere Ziele seines Aufgabenbereiches bedienen kann. Bei zunehmender Zeitdrucklast ist dieses Verhältnis zunehmend variabel. Auf Basis der Ergebnisse werden weiterführende Hypothesen und Erklärungsansätze vorgestellt, die für die Realisierung des beschleunigten Stresstests und somit für die Unterstützung einer Risikoanalyse soziotechnischer Systeme dienen.

1 Wissenschaftliche Zielsetzung und Lösungsansatz

1.1 Einleitung

Die Luftfahrt unserer Zeit ist eine Industrie, die in Europa 18,33 Mio. Flugbewegungen (ACI, 2014) bzw. 9,55 Mio. Flüge (Eurocontrol, 2012) pro Jahr durchzuführen im Stande ist. Im Schnitt waren zwei Unfälle auf 10 Mio. Flüge innerhalb der EASA Mitgliedsstaaten von 2003 bis 2012 zu verzeichnen, womit ein Höchststand der Luftfahrtsicherheit erreicht war. Die Luftfahrt und insbesondere das Luftverkehrsmanagement (ATM) verbinden zahlreiche Subsysteme, bestehend aus Mensch und Technik, zu einem funktionierenden Gesamtablauf. Die Luftfahrt gehört somit auch zu den komplexesten und verteiltsten sicherheitsrelevanten Industrien der Welt (Blom H. , et al., 2001; Brooker, 2010). Der gegenwärtige Stand der Sicherheit ist das Ergebnis einer jahrzehntelangen Entwicklung von standardisiertem Verfahrensdesign und Arbeitsabläufen, die anhand von Unfallanalysen für die bekannten Gefahrensituationen zu einem Optimum ausgelegt wurden. Dies macht ATM robust und widerstandsfähig gegenüber möglichen kritischen Grenzfällen, wobei das menschlich Leistbare unter allen kritischen Betriebsumständen ein sicherheitsrelevanter Aspekt bleibt.

Der insofern weiterhin maßgebliche Beitrag des Menschen zu auftretenden Unfällen und Vorfällen ist eine der Motivationen für die beständige Verbesserung der Luftfahrt. Ein Anteil des Unfallbeitrages zwischen 60 % und 80 % kann heutzutage dem menschlichen Fehler zugeordnet werden (Shappell, et al., 2007; Müller, 2004). Dieser Beitrag zur Unfallentstehung wird kontinuierlich überwacht und ihm wird durch entsprechende risikomindernde Maßnahmen entgegengewirkt. Dies wird durch die obligatorische Berichterstattung über Unfälle und Vorfälle nach ICAO-Standards (ICAO, 2010) und nachfolgenden Gesetzen (FIUUG) sichergestellt, aufgrund derer nationale Aufsichtsbehörden beauftragt sind, Unfallursachen systematisch zur Maßnahmenenergreifung zu ermitteln und zu dokumentieren.

Zusätzlich zu diesem Unfallverhütungsverfahren zielen Modernisierungsprogramme wie das europäische *SESAR* (SESAR, 2012), das U.S.-amerikanische *NextGen* (Federal Aviation Administration, 2014) und das japanische *CARATS* auf eine Erhöhung der Leistungsfähigkeit aller wichtigen Schlüsselbereiche der Luftfahrt ab. Bei SESAR sind diese Umweltverträglichkeit, Kosteneffizienz, Sicherheit und Kapazität (SESAR, 2012). Für die Verbesserung der Sicherheit wird der Faktor 10 bis 2020 genannt. SESAR umfasst die Einführung zahlreicher neuer Systeme in den laufenden Betrieb, wie z.B. das Informationsnetzwerk SWIM¹ und den Flugprofil-zentrierten Flugbetrieb (*trajectory based operations*). Am Beispiel dieser Modernisierungsprogramme wird ersichtlich, dass mit wesentlichen Änderungen der elementaren Abläufe des Flugbetriebes im nächsten Jahrzehnt zu rechnen ist.

Eine signifikante Systemänderung, durch z.B. die Einführung eines neuen Verfahrens oder Instruments an Bord oder am Boden, hat aufgrund der zahlreichen Interaktionen mit

¹ System Wide Information Management (SWIM) ist ein in Entwicklung befindliches Netzwerk zum Austausch von sicherheits-, kapazitäts- und umweltrelevanten Informationen zwischen allen Akteuren des ATM-Netzwerks.

abhängigen Betriebsabläufen neue Gefahrensituationen zur Folge (innovationsinduzierte Gefahren). Häufig wird auch der Begriff „unvorstellbare Gefahren“ verwendet. Um die Sicherheit in der Luftfahrt zu gewährleisten, ist es geltende Annahme, dass das hierdurch entstehende Risiko beherrschbar erscheint, solange die Änderungen überschaubare Bereiche betreffen oder abgestuft in kleinen Schritten vollzogen werden. Geltende Regularien wie z.B. die Verordnung *EC No. 1035/2011* für die Staaten der Europäischen Union verlangen die ordentliche Identifikation von Gefahren, die aus der Systemänderung entstehen (The European Commission, 2011). Die Anwendung angepasster Risikominderungsmaßnahmen soll somit eine Konformität des neuen Systems gegenüber geltenden Sicherheitsstandards gewährleisten. Innerhalb der EU wurde hier ein maximaler Beitrag der Flugsicherungen zum Unfall (*ATM-related Accident*) von $1,0e-8$ pro Flugstunde gemäß (EUROCAE, 2010) definiert. Die Zertifizierung der Flugtüchtigkeit von großen Flugzeugen verlangt für Funktionen, die bei Versagen den Totalverlust des Flugzeugs zu Folge haben können, gemäß Zulassungsvorschrift für große Luftfahrzeuge CS-25 (EASA, 2015) eine maximale Versagenswahrscheinlichkeit von $1,0e-9$ pro Flugstunde.

Die geforderten Zielsicherheiten sind also extrem klein und stellen so eine besondere Herausforderung an ihre Nachweisbarkeit, möglichst vor der Inbetriebnahme, dar. Heutige Methoden wie z.B. der beschleunigte Materialstresstest erlauben es auch für längere Zeiträume, zuverlässige Prognosen für die Lebensdauer technischer Produkte zu ermitteln. Auch hier sind vergleichbar kleine Wahrscheinlichkeiten eines Zielereignisses zu bestimmen. Eine derartige Aussage für Systeme, die den Mensch involvieren, kann mit gegenwärtigen Methoden allerdings nicht geleistet werden. Aus diesem Grund fußen die Verfahren, sowohl bord- als auch bodenseitig, auf standardisierten Arbeitsanweisungen (*Standard Operating Procedures*), die schwerwiegenden Folgen des menschlichen Fehlers durch vordefinierte Arbeitsschritte einzugrenzen oder gar vorzubeugen versuchen. Trotz aller risikomindernden Maßnahmen zeigen aktuelle Statistiken des menschlichen Fehlers die unverminderte Relevanz für gegenwärtige Unfälle. Aufgrund des einmaligen Charakters der Kombination von Faktoren, die zur Unfallentstehung beigetragen haben, kann mit Hilfe eines standardisierten Verfahrensdesigns allein kein vollständiger, vorbeugender Schutz gewährleistet werden. Sorglosigkeit oder Überforderung können hier exemplarisch genannt werden, welche die Grenzen des leistbaren Schutzes durch den Menschen begründen. Die Herausforderung besteht dann darin, das Risiko eines Fehlerverhaltens des Menschen in einer Wechselwirkung mit der Technik und der Umgebung für den Flugbetrieb abzuschätzen.

Eine besondere Herausforderung stellen dabei fundamentale Änderungen in Technik und Verfahren dar, die sogenannte Paradigmenwechsel in der Luftfahrt einleiten. Als Beispiel kann die *Remote Tower*-Technologie in der Flugplatzkontrolle genannt werden, die die Arbeit des Platzverkehrslotsen mit Hilfe einer videobasierten Außensicht unterstützen (Abbildung 1). Die dadurch gewonnene Ortsunabhängigkeit der Flugplatzkontrolle soll es nach erfolgtem Sicherheitsnachweis der Flugsicherungsdienste gestatten, die Kontrollleitstellen mehrerer Flugplätze an einem Ort zu konzentrieren. Die Interaktion des Remote Tower zwischen Außensicht, Instrumenten und dem Menschen kennzeichnen ein repräsentatives *soziotechnisches System*, das nach Definition (Emery & Trist, 1960) Technologie, Menschhandlungen und die Betriebsumgebung eng miteinander verknüpft.

Derartige Systeme stellen mit den dazugehörigen neuen Verfahren eine Systemänderung dar, an denen bisher praktizierte Methoden zur Risikobestimmung und -bewertung an die Grenzen stoßen, das Risiko bereits vor der Inbetriebnahme nachvollziehbar und vorhersagestark einschätzen zu können (Brooker, 2010). Bei dem aktuellen Zulassungsverfahren des Remote Tower Sundsvall kann dieses Problem an dem Beispiel der Gefahr eines plötzlichen Ausfalls eines oder mehrerer von den insgesamt 14 Monitoren verdeutlicht werden. Zwei Luftfahrzeuge könnten dann durch den Platzverkehrslotsen nicht mehr beobachtet werden. Für die bei dem Ausfallszenario in Erwägung gezogenen hypothetischen Betriebssituationen konnten belastbare Aussagen zum Risiko bis heute nicht ermittelt werden.

Diese Methoden werden im Folgenden unter dem Begriff *Risikoanalyse* zusammengefasst, die die Bestimmung des Risikos eines Systems in Verbindung mit den Ursachen zum Ziel hat. Der später folgende Abschnitt 1.3 beschreibt das Ziel und den methodischen Hintergrund hierzu ausführlicher.



Abbildung 1: LFV- Remote Tower Center in Sundsvall

Die soeben erwähnte Grenze betrifft vornehmlich Methoden, die sich grundlegend auf Expertenaussagen stützen und hypothetische Gefahrensituationen identifizieren und hinsichtlich ihres potentiellen Schweregrades bewerten (de Jong, 2004). Diese Methoden sind vorteilhaft, da anhand des betrieblichen Erfahrungsschatzes von Experten wie z.B. Fluglotsen oder Piloten eine erste Abschätzung über den Einfluss eines neuen Systems auf das Gesamtverhalten gewonnen werden kann. Jedoch können auch Experten, insbesondere in Hinblick auf die Komplexität des Flugbetriebes, nicht alle Auswirkungen eines neuen Systems zuverlässig vorhersagen und beurteilen. Auch ist es so, dass die Vorstellung und vorherrschende Prämissen der Arbeitsausführung von Experten sich stark an dem Etablierten und Vorherrschenden orientieren. Sie versuchen konventionelle Denkmuster auf neuartige Arbeitsumgebungen zu projizieren. Die Erwartungshaltung und das Unterstellen bzw. Anwenden bekannter Muster durch den Experten lässt erahnen, dass eine derartige Gefahrenidentifikation keine hinreichende Beobachtungsbasis aufweisen kann. Innovationsinduzierte bzw. unvorstellbare Gefahren können aufgrund ihrer Neuheit keinem vorab bekannten Muster entsprechen. Die Wahl eines Ansatzes, der die Gefahrenidentifikation auf Basis von beobachteten Einzelfällen des Betriebs vornimmt, wäre wohl vorzuziehen. Eine solche Ergründung von möglichen Gefahren aus dem Beobachteten folgt einer sogenannten induktiven Logik der Schlussfolgerung, die allerdings die Inbetriebnahme des neuartigen Systems voraussetzt.

Der Nachteil der expertenbasierten Risikoanalyse kommt z.B. im Rahmen der *Deployment*-Phase von SESAR ab dem Jahre 2015 zum Tragen, die zahlreiche neuartige soziotechnische Systeme in den Flugbetrieb einbinden wird. Aufgrund der Größenordnung dieser Änderungen ist eine vorab gewinnbare Einschätzung des resultierenden Risikos als unrealistisch zu erachten (Brooker, 2010). Die Nutzung von expertengestützten Schätzungen ist dennoch üblich, da belastbare Daten aus empirischen Untersuchungen häufig nicht verfügbar sind (Mosleh, B., & Apostolakis, 1988). Folglich sind die geleistete Gefahrenidentifikation und die dazugehörige Schätzung von Unfallwahrscheinlichkeiten zum Beitrag eines Unfalls nicht garantiert vollständig:

„If a human error is not identified then it does not matter how accurate the quantification technique is, since the error will simply not appear in the risk assessment” (Kirwan, 1992, S. 302).

Die zuverlässige Identifikation aller oder zumindest der kritischsten Fälle entpuppt sich somit als eine methodische Herausforderung, für die bis heute keine zufriedenstellende Lösung gefunden werden konnte. Die Suche nach den relevantesten Gefahren (Risikoschwerpunkte) und den dazugehörigen Ursachen eines neuen Systems ist auch heute nach wie vor ein zeitraubender iterativer Prozess, der zu einer Verlangsamung des Entwicklungsprozesses soziotechnischer Systeme führen kann (Brooker, 2010). Die Gefahrenidentifikation industrieller Entwicklungsprojekte begrenzt sich aus diesem Grund auf Gefahren, die sich auf dem zu dem Zeitpunkt bekannten Wissen begründen lassen und dadurch in einem akzeptablen Rahmen der geltenden Zulassungsverfahren neuartiger Systeme liegen können.

In den letzten 15 Jahren hat sich die modellierungsgestützte Risikoanalyse zum probaten Hilfsmittel zur Gewinnung quantifizierter Risikowerte mit großem Potential entwickelt. Hierbei können z.B. das Projekt TOPAZ (Blom H. , et al., 2001) oder die Simulation auf Basis von modellbasierten Fluglotsenagenten (Vogel, Thiel, & Fricke, 2012) genannt werden, die das Wirken von Einzeloperatoren auf den Flugbetrieb anhand von softwarebasierten Agenten mit einzustellenden Verhaltensparametern modellieren und simulieren. Ein wesentlicher Schwachpunkt bleibt allerdings die Gültigkeit der Modellierung von menschlichen Fehlern in der Verkettung zum Unfall:

„Errors are likely to be made when designers apply error modeling techniques” (Johnson, 1999, S. 517).

Dieser Fall tritt ein, wenn menschliches Verhalten anhand von entsprechenden Faktoren nicht beschrieben und Kausalrelationen sowie Individualverhalten nicht im genügenden Maße berücksichtigt werden können, da z.B. empirische Daten nicht im notwendigen Maße verfügbar sind. Ein ebenso verantwortlicher Umstand ist die Komplexität des menschlichen Gesamtverhaltens im Betrieb, deren Einzelparameter hinsichtlich ihrer Relevanz eingeschätzt oder aufwendig experimentell bestimmt werden müssen. Es entsteht in Folge eine enorme Anforderung an die Modellierung des Individualverhaltens und an die dazugehörigen Varianzen, um das Risiko nach realen Maßstäben abschätzen zu können. Dieser Aufwand ist erforderlich, obwohl die Entscheidungsprozesse des Operators im Flugbetrieb bereits eine Verfahrensstandardisierung („Prozeduralisierung“) gemäß *ICAO PANS-ATM Doc. 4444* bzw. das *ICAO Annex 3 – Rules of the air* sowie in Deutschland rechtsverbindlich durch die

„Betriebsanweisung Flugverkehrsdienste“ (BA-FVD) der Deutschen Flugsicherung erfahren haben (DFS, 2011).

Im Kern scheint die Schwierigkeit gegenwärtiger Risikoanalysen in dem notwendigen A-priori-Wissen über mögliche Unfallursachen zu bestehen, deren Risikobeitrag einzeln und zumeist prospektiv bestimmt werden muss. Hierbei besteht aufgrund des begrenzten Wissens über mögliche Ursachen grundsätzlich das Problem, die wichtigsten für die Bestimmung des Risikos nicht mit einbezogen zu haben. Dies ist insbesondere dann von Nachteil, wenn das breite Ursachenfeld eines neuartigen soziotechnischen Systems die vermeintlich vollständige Nachweisprüfung jedes einzelnen potentiellen sicherheitsrelevanten Faktors nicht unter Aufwendung realistischer Mittel zulässt. Für diesen Einsatzzweck ist der prospektive Untersuchungsansatz nicht geeignet. Dieser Schwachpunkt wird insbesondere innerhalb der Entwicklung des Konzeptes und des Designs eines soziotechnischen Systems sichtbar, wenn Risikoschwerpunkte effektiv durch Sicherheitsanforderungen kontrolliert oder gar eliminiert werden sollen. Eine flächendeckende Identifikation von Risikoschwerpunkten in den möglichen Ursachen ist insbesondere zu diesen Phasen unabdingbar.

Es scheint aufgrund dieses inhärenten Merkmals des prospektiven Untersuchungsansatzes notwendig, Methoden der Risikoanalyse zu entwickeln, die von A-priori-Annahmen bzw. Theorien über reale Gefahren- und Ursachendefinition befreit oder zumindest weitestgehend entlastet sind. Dieses trifft auf Risikoanalysen mit dem retrospektiven Untersuchungsmerkmal (*ex post facto research*) zu (Kerlinger, 1964), die lediglich den Unfall selbst als vorab bekanntes Ereignis voraussetzen. Eine derart geführte Bestimmung des Risikos unterstellt keine initialen Gefahrenereignisse und keine unabhängige Variable, deren kausale Relation zum Unfall es zu beweisen gilt. Das ermittelte Risiko bedingt also keines Vorwissens und gestattet die Verifikation des Risikos hinsichtlich der Erfüllung einer akzeptablen Zielsicherheit. Dieses Anwendungsszenario stellt die ideale Lösung des Problems notwendigen Vorwissens innerhalb dieser Arbeit dar.

Der Nachteil einer derartig gerichteten Untersuchung ist allerdings die fehlende Möglichkeit zur Verifikation von Kausalhypothesen, da andere Erklärungsmöglichkeiten außer mittels Beobachtung mangels Kontrolle nicht ausgeschlossen werden können (Kerlinger, 1964). Da das Luftfahrtsystem wie dargelegt nach weitestgehend festen prozeduralen und standardisierten Ablaufmustern organisiert ist, können die Freiheitsgrade der möglichen Erklärungen nach Datenlage und Plausibilität eingegrenzt werden. Auf Grundlage dieses Wissen und von Expertenaussagen können Unfallentstehungen anhand von sogenannten Fehlerbaummodellen ergründet werden. Diese können die Ursachenmöglichkeiten in der Technik, dem Menschen, den Verfahren und Organisation sodann eingrenzen und geeignete Ursachenfaktoren ergründen. Eine derartig konzipierte Methode zur Bestimmung des Risikos könnte somit zu einer beschleunigten Entwicklung von Sicherheitsanforderungen für sicherheitsrelevante Funktionen eines soziotechnischen Systems beitragen, da Risikoschwerpunkte effektiv eingegrenzt werden können. Die Anwendung eines derartigen Ansatzes verspricht einen enormen Bedeutungsgewinn der Risikoanalyse für die Entwicklung soziotechnischer Systemen in der Luftfahrt.

Auf Grundlage der soeben geführten Argumentation ist es Motivation dieser Arbeit, einen Beitrag zur Entwicklung retrospektiver Methoden der Risikoanalyse zu leisten. Der hierfür gewählte Ansatz fällt in den Bereich der experimentellen Untersuchungsmethoden, die

die Bestimmung des Risikos und der Ursachen auf Daten von Echtzeitsimulationen (eng. Human-In-The-Loop Simulation) stützt. Die Echtzeitsimulation verspricht dabei die Berücksichtigung komplexer Interaktionsvorgänge zwischen dem Operateur und dem betrieblichen Umfeld auf Basis von Beobachtungen in einer kontrollierten Arbeitsumgebung. Die Untersuchung der möglichen erzielbaren Qualität belastbarer Daten und die dazugehörigen notwendigen Bedingungen für eine effektive Unterstützung neuartiger Systeme ist somit die zentrale Motivation.

Der Effekt der Rechtzensur wird als das entscheidende Problem der Methode der retrospektiven Risikoanalyse identifiziert und diskutiert. Eine zur Lösung der Rechtzensur entwickelte Methode stützt sich auf ein zeitdruckbeschleunigendes Stresstestverfahren, welches das Problem der de facto stets begrenzt zur Verfügung stehenden Simulationszeit lösen soll. Der Effekt der Rechtzensur wird im Abschnitt 1.4.3.2 näher vorgestellt. Das erklärte Ziel dieser Methode ist die Verifikation einer geforderten Zielsicherheit² für lange Zeiträume auf Basis kürzerer Versuchszeiträume. Die Wirksamkeit eines eigens dazu entwickelten Zeitdruckinduktionsverfahrens wird im Rahmen dieser Arbeit mit Hilfe einer experimentellen Pilotstudie getestet. Anhand von drei Probanden wird dieses neuartige Verfahren erprobt und Untersuchungshypothesen für weiterführende Studien werden entwickelt. Die Arbeit soll schließlich einen Erkenntnisbeitrag zur zuverlässigen und reproduzierbaren Gewissheit über den sicheren Betrieb neuartiger soziotechnischer Systeme liefern. Die Abhängigkeit von expertenbasierten Schätzungen soll somit verringert werden.

Kapitel 1 erläutert Struktur und Arbeitsschritte zur Beschreibung des Vorgehens und zum Ziele dieser Arbeit. Zentrales Ergebnis dieses Kapitels ist die Herleitung der Untersuchungshypothese im Abschnitt 1.5.2. Die Struktur der Arbeit bildet sodann den letzten Abschnitt dieses Kapitels. Für die Herleitung der Vorgehensweise beinhaltet es zusätzliche Abschnitte, die dem Verständnis des zu untersuchenden Problems dienen mögen und dazu notwendige Zusammenhänge vorstellen. Zunächst geben Statistiken des menschlichen Fehlers in der Unfallforschung einen allgemeinen Überblick über die Relevanz dieses Fehlers. Die Problematik der Nutzbarkeit von Analysen des menschlichen Fehlers für die Risikoanalyse wird herausgestellt. Anschließend werden Ansätze gegenwärtig praktizierter Methoden der Risikoanalyse vorgestellt, die je nach Ziel und Zweck des zu bestimmenden Risikos unterschiedlichen Konzepten und dazugehörigen Prämissen folgen. Zur Strukturierung nach charakteristischen Merkmalen der unterschiedlichen Methoden wird eine Klassifikation vorgestellt, die nach der jeweiligen Logik der Untersuchung unterscheidet. Anschließend erfolgt eine umfängliche Analyse des Wertes von Echtzeitsimulatoren und deren Nutzen für die Risikoanalyse. Ein wesentlicher Punkt sind die Grenzen der Nutzbarkeit dieses Mittels und die bislang ungelösten Ursachen hierfür, die vornehmlich in der begrenzten Simulationszeit und somit begrenzten Beobachtbarkeit von sicherheitsrelevanten Ereignissen liegen. Auch spielt hierbei eine Rolle, dass Simulatoren zur Herstellung der Arbeitsumgebung auf physikalischen bzw. technischen Modellen aufsetzen, die einen wesentlichen, aber stets auch nur begrenzten Ausschnitt des realen Flugbetriebes liefern können.

² Eine Zielsicherheit ist eine maximal akzeptierte Unfallwahrscheinlichkeit (Eurocontrol, 2001). Eine nähere Erläuterung erfolgt in Abschnitt 2.1.2.

1.2 Der menschliche Fehler in der Luftfahrt

1.2.1 Der statistische Unfallbeitrag

Bei Betrachtung der historischen Entwicklung des Anteils menschverursachter Unfälle pro Flugstunde kann ein sinkender Trend festgestellt werden, der dem fallenden Trend der technisch-mechanischen Ursachen folgt (Shapell & Wiegmann, 1996). Die Abbildung 2 zeigt diese Entwicklung von 1977 bis 1992 der Klasse A, B, und C Unfall- und Vorfaltrate³ beispielhaft anhand ausgewerteter Unfall- und Vorfallberichte der U.S.-Marineluftfahrt. Die gleichzeitig fallenden Trends zeigen dabei auf, dass sich die zeitlichen Entwicklungen über die Jahre kongruent verhalten. Eine Erklärung dafür kann sein, dass zumeist auf einen technischen Fehler ein menschlicher Fehler folgt. Die enge Verbindung von Mensch und Maschine verdeutlicht das Funktionieren des Flugbetriebes als soziotechnisches System.

Bei Auswertung der relativen Anteile innerhalb der Ursachen konnte der menschliche Beitrag für das Jahr 1996 auf zwischen 60 und 80 Prozent (Shappell, et al., 2007) bzw. 75 % (Müller, 2004) beziffert werden. Der große relative Anteil ist dabei aber auch die Folge des kontinuierlich sinkenden Anteils an technisch-mechanischen Ursachen, der den Anteil an menschfehlerbedingten Unfällen stärker gewichtet.

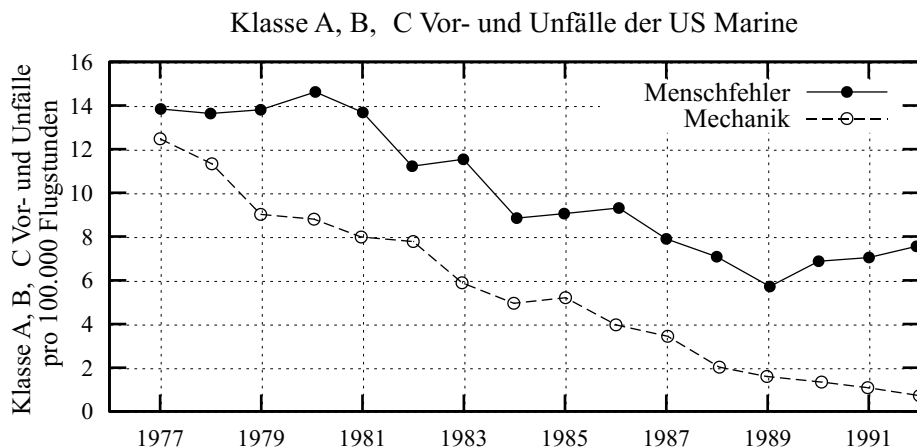


Abbildung 2: Vor- und Unfallraten der U.S. Marineluftfahrt von 1977–1992 (Shapell & Wiegmann, 1996)

Die Verursachung eines Unfalls durch menschliche Fehler ist anhand der Erfahrung dokumentierter Unfälle erklärbar. Bei der Untersuchung von Unfallursachen konnten die folgenden Regelmäßigkeiten festgestellt werden:

„It is generally accepted by accident investigators in the field that aviation accidents are the result of a chain of events culminating with the unsafe acts of operators (i.e., aircrew)” (Shappell, et al., 2007, S. 228),

„By far the most frequent safety-critical situation (37.8 % of all events) consists of the following combination of risk factors: 1. A complication develops. 2. In this situation of increased stress a human error occurs. 3. The negative effects of the error cannot be

³ Klasse A, B und C: Jegliches Schadensereignis an Flugzeug und Besatzung mit einem Wert von wenigstens 50.000 US-Dollar oder/und Verletzungen am Menschen bzw. Todesfall (US Navy, 2016).

corrected or eased because there are deficiencies in team interaction on the flight deck” (Müller, 2004, S. 559).

„This chain starts with the first precursor leading to the outcome (e.g., safety relevant event, incident, or accident). In this respect, the analysis of contextual factors should cover the entire chain and assess the relevant context for each link in the chain.” (Subotic, Ochieng, & Sträter, 2007, S. 859).

Die Wirkung fehlerhafter menschlicher Tätigkeiten in einer Gesamtverkettung der Faktoren ist demnach ein Kernelement der Unfallentstehung. Im Bereich der Reaktorsicherheit wird die grundsätzliche Menschfehlerwahrscheinlichkeit nach empirischen Studien auf wenigstens 0,1 % pro Einzelhandlung (z.B. Bedienaktion) geschätzt (Swain & Guttman, 1983). Ebenso übertritt der Mensch in der Luftfahrt mutmaßlich häufig zulässige Grenzen, was allerdings ohne ernsthafte bzw. dokumentierte Auswirkungen bleibt. Aus diesem Grund offenbaren dokumentierte Unfälle lediglich einen Teil der täglich auftretenden menschlichen Fehler. Jedoch sind menschliche Fehler differenzierter zu betrachten, wenn man ihren Beitrag zur Unfallentstehung verstehen will. Eine Betrachtung der ursächlichen Fehlerarten anhand von ausgewählten Unfallereignissen zeigt beispielhaft die unterschiedlichen Merkmale, die ein Auftreten haben kann.

- Unfälle wie der des Fluges Air France AF 447 von Rio de Janeiro nach Paris im Jahre 2009 zeigen beispielhaft, welche tragischen Folgen eine inkorrekte Einschätzung der vorliegenden Gefahrensituation und die daraus resultierende fehlerhafte Führung des ansonsten hochautomatisierten Airbus A330 haben kann. Dieser sogenannte *Decision Error* kann nach dem *Human Factors Analysis and Classification System* (HFACS) als Planungs- oder Denkfehler bzw. Entscheidungsfehler erachtet werden (Wiegmann & Shappell, 2001), in der vorliegende Informationen wie z.B. das Abschalten des Autopiloten und die widersprüchlichen Geschwindigkeitsangaben nicht zu einem schlüssigen Gesamtbild verbunden und somit unangemessene Entscheidungen getroffen wurden (BEA, 2012). Allerdings haben in diesem Fall ebenso Fehler zum Unfall beigetragen, die nicht diesem Typus menschlicher Fehler zugeordnet werden können, wie z.B. die emotionale Überforderung der Piloten.
- Dem gegenüber steht der *Skill-based Error*, der einen zumeist unbewussten Fehler des Operateurs beschreibt, der durch „handwerkliche“ Fehler hervorgerufen wird (z.B. fehlerhafte Wahrnehmung, unsorgfältige Bedienung). Als Beispiel für einen menschlichen Fehler dieses Typs kann der Unfall des Asiana Airlines Flugs 214 in San Francisco im Jahre 2013 genannt werden. Durch die fehlerhafte Annahme des Piloten, die Autothrottle-Funktion aktiviert zu haben, wurde die Mindestgeschwindigkeit der Boeing 777 im Endanflug unterschritten.

Sowohl der Entscheidungsfehler (39,2 %) als auch der fertigkeitbasierte Fehler (42,5 %) stellen nach Analyse von NTSB-Daten aus dem Zeitraum von 13 Jahren (1990 bis 2002) mit einer Stichprobenanzahl von 181 Unfällen in der kommerziellen Luftfahrt den Großteil der Fehler dar (Shappell, et al., 2007). Die übrigen Anteile fallen dem Wahrnehmungsfehler (5,5 %) und der *Violation* als Form des Regelverstoßes zu (17,1 %). Fehlende Sorgfalt und Denkfehler des Menschen scheinen hiernach entscheidende sensitive Ursachenfaktoren des

Unfalls zu bilden. Der zeitliche Trend der relativen Anteile dieser vier Fehlertypen ist in Abbildung 3 illustriert und verdeutlicht die unveränderte Präsenz dieser Schwerpunkte gegenüber den anderen beiden Fehlertypen.

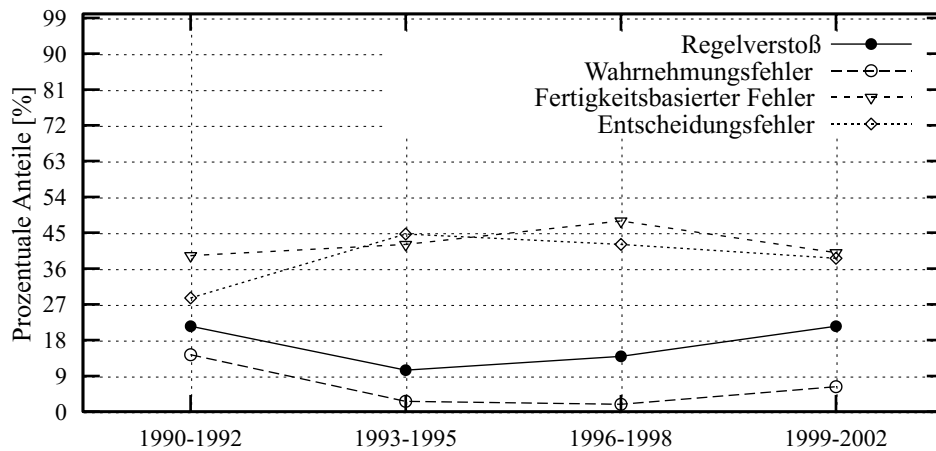


Abbildung 3: Ursachenanteile des menschlichen Fehlers in der kommerziellen Luftfahrt, klassifiziert nach HFACS (Shappell, et al., 2007)

Bei der weiteren Zurückverfolgung dieser Fehlertypen spielen elementare Faktoren eine Rolle, die die Wahrnehmungsprozesse des Menschen negativ beeinflussen. Diese Faktoren sind auch bekannt als *Performance Shaping Factors*, die im Wesentlichen Stressfaktoren darstellen (Swain & Guttman, 1983). Der Stressfaktor produziert demnach allgemein eine Spannung, die einen Handlungszwang erzeugt. Beispielhaft führt (Green, 1985) drei wesentliche Stressfaktoren für die Minderung der Leistungsfähigkeit von Piloten auf: Umgebungsbedingungen, akute stressbedingte Reaktionen bei Extrembedingungen und Lebensbedingungen. Die weitere Suche nach Faktoren kann in einer unerschöpflichen und extrem vielschichtigen sowie detaillierten Auflistung enden, die für eine Anwendung im Bereich der Designevaluation häufig wenig hilfreich ist. Eine anwendungsspezifische und zugeschnittene Definition von derartigen Faktoren ist aus diesem Grund einer allgemeingültigen Definition vorzuziehen.

Im Bereich der Reaktorsicherheit wurde geschätzt, dass die Wahrscheinlichkeit menschlichen Fehlverhaltens um den Faktor 10 für Novizen und den Faktor 5 für Experten steigt, wenn der Mensch eine Gefahrensituation mit einer starken emotionalen Stresserregung (*Thread Stress*) erfährt (Swain & Guttman, 1983). Eine wichtige Randbedingung stellt die Gesamtheit der Performance Shaping Factors dar, wie z.B. emotionale Stabilität, Ausbildung und die allgemeine Verfassung (Bubb, 2005).

Das Wissen über das Vorliegen von Fehlertypen, gleich welches Klassifizierungssystem auch verwendet wird, kann aufgrund des nicht näher spezifizierten Erscheinungsmusters nicht zu einer effektiven Risikominderung verwendet werden. Hierauf beruht die Hauptkritik an den Klassifizierungssystemen und den Methoden zur Bestimmung der jeweiligen Anteile. So hat die A-priori-Klassifizierung des menschlichen Fehlers bislang wenig praktischen Einfluss auf das Design von neuartigen Systemen gehabt (Johnson, 1999). Gründe hierfür sind z.B., dass die Kenntnis über die Existenz sowie die Relevanz und die beobachtete Häufigkeit eines spezifischen Typus von menschlichem Fehler ohne

Informationen über das Wirken in einem situativen Gesamtbild (Kontextinformationen) keinen Rückschluss auf eine ggf. notwendige Designänderung zur Minderung der Eintrittshäufigkeit liefern. Der auf den Menschen einwirkende Stress kann allerdings aufgrund der damit verbundenen Leistungsminderung als gemeinsamer „Ursachennenner“ des Risikobeitrages durch den Menschen identifiziert werden.

1.2.2 Modelle des menschlichen Unfallbeitrags

Da Unfallanalysen die Relevanz des Zusammenwirkens von Umgebung, Technik und Mensch für das entstehende Risiko verdeutlichen, kann die dahinterliegende Systematik anhand von Theorien erklärt werden. Der Auftritt des menschlichen Fehlers kann allgemein durch eine Abweichung von der tolerierbaren Arbeitsqualität erklärt werden, bei der eine Mindestqualität in Gegenwart einer vorliegenden sicherheitsrelevanten Situation unterschritten wird und eine Eskalation zugunsten eines Unfalls entsteht (Bubb, 2005). Diese allgemeine Theorie stützt sich dabei auf eine durch die Betriebssituation geforderte Mindestqualität, anhand derer die Qualität menschlicher Tätigkeiten in „akzeptabel“ und „nicht akzeptabel“ eingeteilt werden kann. So ist es folglich möglich, dass im Einzelfall die gleiche Tätigkeit je nach Toleranz der Situation einen Unfall zur Folge hat oder auch nicht und das Ergebnis somit über das Vorliegen eines menschlichen Fehlers im Nachhinein entscheidet. Im Sinne dieser einfachen Theorie erhöht zunehmender Stress die Wahrscheinlichkeit für das Unterschreiten einer durch die Betriebssituation geforderten Mindestqualität.

Dieser Ansatz des Zusammenwirkens wird durch das Modell *Trajectory of Accident Opportunity* (auch bekannt als *Swiss Cheese*-Modell) beschrieben (Reason, 1990). Dieses Modell hat ihren initialen Ausgangspunkt während des Systemdesigns bei den *Fallable Decisions* durch die Systemdesigner und höheren Manager-Entscheidungen. Diese sogenannten *Latent Errors* begründen sich in Schwächen in der Auslegung der Verfahren und der Technik, die lange verborgen bleiben können. Unfallanalysen kommen aus diesem Grund häufig zu dem Schluss, dass der Fehler darauf gewartet hat, einen Unfall erzeugen zu können (Rasmussen, 1997). Zum Tragen kommen diese Unzulänglichkeiten bei Eintritt eines *Active Errors*, bei dem der Operateur des Systems eine sofortige negative Wirkung erhält. Das darauf basierende Modell stellt den Eintritt des Unfalls nach Gefahrenauftritt als Abhängige von den „Psychologischen Vorbedingungen“, dem sogenannten *Unsafe Act*, und dem Versagen von Schutzbarrieren dar. Der *Unsafe Act* ist dabei ein menschlicher Fehler, bei dem die Menschlichkeit eine Mindestqualität letztendlich unterschreitet. Er wird durch Reason wie folgt beschrieben:

„the commission of unsafe acts is determined by a complex interaction between intrinsic system influences and those arising from the outside world. This has to do both with protean environmental factors and with the particular form of the existing hazard” (Reason, 1990, S. 206).

Es fasst den Beitrag von menschlichen Fehlern, Organisation und Technik in der Interaktion zusammen und ist innerhalb der *Human Factors* und *Human Reliability*-Wissenschaften weitestgehend akzeptiert. Reasons Modell der Unfallentstehung stieß auf Kritik, da die Modellierung der Kausalfaktoren wenig praktischen Nutzen für die zukünftige Unfallverhütung bietet (Luxhøj & Kauffeld, 2003). Der Ansatz, die Organisation als Bestandteil der Unfallentstehung zu erfassen, lenkt den Fokus der Unfallforscher auf *Latent*

Errors, die jedoch unzureichenden Rückschluss darauf zulassen, welche *Active Errors* am Ende entscheidend sind (Shorrock S. , 2005).

Reason verdeutlicht durch sein Modell, dass jedweder menschliche Fehler für die Unfallentstehung spezifisch zu den vorliegenden Betriebsbedingungen von Technik, Umwelt und Organisationen passen muss. Es unterstreicht die Schlüsselrolle der Arbeitsumgebung des Menschen, die eine Eskalation der Flugbetriebsicherheit bei Eintritt eines spezifischen menschlichen Fehlers aufgrund von möglichen Latent Errors begünstigen kann. Aufgrund der grundsätzlichen Fehlbarkeit des Menschen verdeutlichten diese Ausführungen die entscheidende Bedeutung der Entwicklung einer fehlertoleranten Arbeitsumgebung des Menschen.

1.3 Die Probleme gegenwärtiger Methoden der Risikoanalyse

Die Risikoanalyse ordnet sich methodisch als Arbeitsschritt einer Risikobewertung ein. Was als Risiko verstanden werden soll, ist dem jeweiligen Anwendungskontext angepasst geeignet zu definieren. Im übergeordneten Sinne handelt es sich hier um die Unsicherheit, ein Ziel zu erreichen:

„the effect of uncertainty on objectives” (Purdy, 2010, S. 882).

Insbesondere auch in der Luftfahrt ist es oberstes Ziel, Schaden an Mensch und Technik zu vermeiden. Die Organisation ICAO erachtet dazu passend das Unfallereignis und seine Vorfalleignisse als oberste Schweregrade einer Konsequenz und setzt sie somit zur absoluten Maßgröße des Risikos der operationellen Sicherheit:

„...the elimination of aircraft accidents and/or serious incidents remains the ultimate goal...” (ICAO, 2013c, S. 2-1).

Somit stellt die Wahrscheinlichkeit für den Eintritt eines Unfalls die Definition für das betriebliche Risiko im Flugbetrieb dar. Das Ziel der Risikoanalyse ist demnach die Bestimmung der Unfallwahrscheinlichkeit aus Gefahrensituationen, die eine Zusammensetzung von Ursachenfaktoren aus angewandten Verfahren, Training, Technik, Arbeitsinstrumenten oder anderen bilden können (ICAO, 2013c). Das Gefahrenereignis ist dabei wie folgt definiert:

„‘Gefahr‘ ist jede Bedingung, Vorfall oder Umstand, die einen Unfall verursachen könnte” (The European Commission, 2011, S. 271/25).

Dies schließt die Verursachung von sicherheitsrelevanten Vorfällen mit ein, die in einem kausalen Zusammenhang mit dem Unfalleintritt stehen. Die Gefahr kann dabei eine umfassende Beschreibung von situativen Faktoren, wie z.B. Wetter und Verkehrssituation, umfassen und sowohl bestimmte Menschfehler oder andere Fehler, wie z.B. Hardwarefehler, unterstellen. Wesentliches Merkmal ist dabei, dass der Eintritt der Gefahr ein Glied in der Verkettung der Ursachen zum Unfall bildet. Charakteristische Untersuchungsergebnisse der Risikoanalyse sind

1. die Wahrscheinlichkeit des Schadens nach Eintritt der Gefahr und
2. die Bestimmung des dazugehörigen Schweregrades.

Zusätzlich kann zu der Analyse

3. die Bestimmung der Risikosensitivität⁴ der einzelnen Parameter gehören, aus denen sich das Gefahrenereignis zusammensetzt. Zu diesen gehören situative Parameter, wie z.B. die verbliebene Menge Treibstoff in einem Flugzeug, die bei erforderlichem Anfliegen eines Ausweichflugplatzes (*Destination Alternate Aerodrome*⁵) ein maßgebender Faktor für die Zeitdringlichkeit der Landung ist. Auch wenn es sich hierbei um ein Standardverfahren handelt, stellt es eine ungewünschte Situation für die Besatzung dar, die Stress verursachen kann.

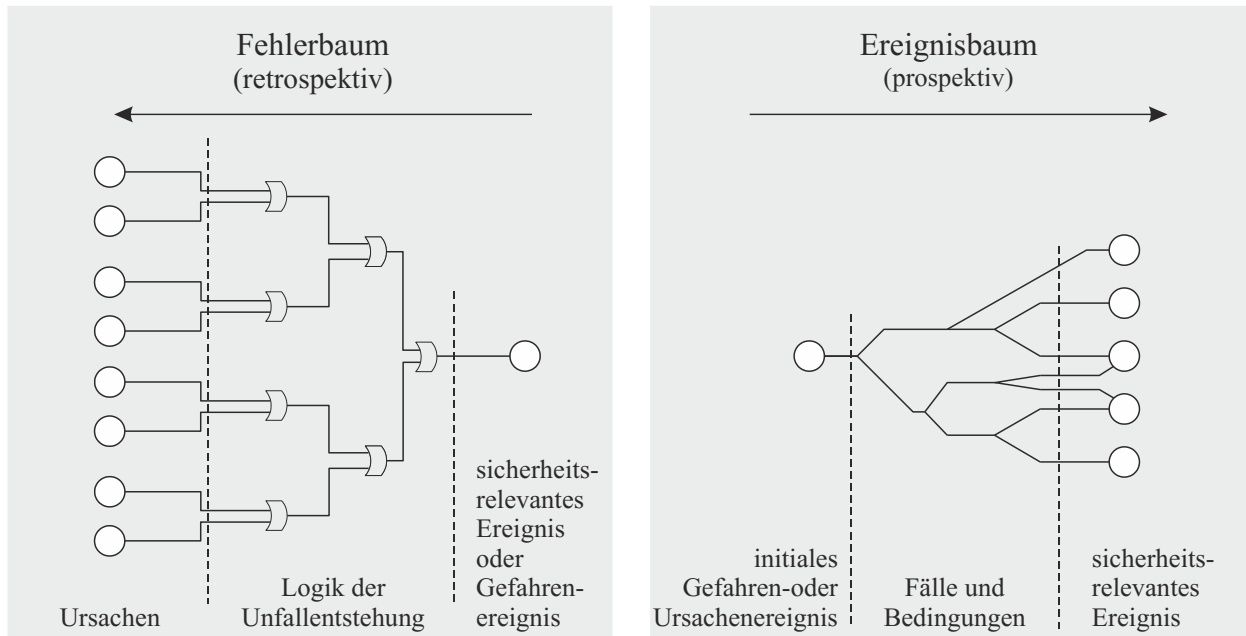


Abbildung 4: Typische Modelle zur Abbildung von Relationen zwischen Ursachen und dem Unfallereignis

Neben der soeben vorgestellten „klassischen“ Analyse ist es bei der Wahl einer geeigneten Methode der Risikoanalyse allgemein hin freigestellt, jedwedes Modell zu verwenden, das eine Aussage zu einem Zusammenhang zwischen einem Unfallereignis bzw. einem Vorfall wie z.B. die Verletzung von Staffelungsminima oder andere meldepflichtige Ereignisse im Sinne der Luftverkehrs-Ordnung (LuftVO) § 9 und seinen Ursachen zu treffen vermag. Aufgrund der Gemeinsamkeit des Un- und Vorfalls, in Verbindung mit der jeweiligen Wahrscheinlichkeit ein Risiko bestimmen zu können, werden beide Ereignistypen als sogenanntes *sicherheitsrelevantes Ereignis* begrifflich zusammengefasst. Die Abbildung 4 stellt hierzu zwei typische Modelle vor, die in der Sicherheitsbewertung zur Darstellung der Kausalrelation benutzt werden. Die Untersuchungsrichtung ist im jeweiligen Feld mit einem Pfeil gekennzeichnet. Wesentliche Unterschiede zwischen den Modellen bestehen in der Untersuchungsrichtung der Analyse und der Nutzung deterministischer Logik des Fehlerbaums. Ihm steht die probabilistische Modellierung des Ereignisbaums gegenüber, der nach möglichen Fällen und den dazugehörigen Wahrscheinlichkeiten der kontextuellen

⁴ Die Risikosensitivität ist ein Übertragungsfaktor, der die Änderung des Risikos bei Variation von situativen Faktoren oder anderen Umgebungsfaktoren der Gefahr ausdrückt.

⁵ Der *Destination Alternate Aerodrome* ist gemäß *ICAO Annex 6* ein Flughafen, den ein Flugzeug für den Fall anfliegt, dass der geplante Zielflughafen für eine Landung nicht zur Verfügung steht.

Bedingungen unterscheidet (bedingte Konsequenzen). Bei Unterstellung von Wahrscheinlichkeiten für die Ursachenzustände kann der Fehlerbaum ebenfalls für eine probabilistische Analyse verwendet werden.

Um die Probleme gegenwärtiger Methoden der Risikoanalyse zu verstehen, werden im Folgenden fünf Methoden exemplarisch vorgestellt. Die Schlussfolgerung liefert eine Erklärung für die systematischen Defizite gegenwärtiger Methoden für die Erbringung des Sicherheitsnachweises neuartiger soziotechnischer Systeme. Das Untersuchungsziel dieser Arbeit kann auf Basis dieser Erklärung weiter eingegrenzt werden.

1.3.1 Die „Safety Assessment Methodology“

Diese kurz SAM genannte Methode umfasst die Analyse des Risikos als Teil eines *Functional Hazard Assessments* (FHA), beginnend beim Auftreten eines als initial festgelegten Gefahrenereignisses. Die Risikoanalyse wird auf Grundlage eines Ereignisbaummodells durchgeführt, das identifizierte Gefahren als Schlüsselereignisse darstellt und von denen ein Risiko bezüglich des Betriebes ausgeht. Das aus dem Gefahrenereignis entstehende Risiko wird dabei mit Hilfe von probabilistischen Fallunterscheidungen modelliert, welche je einer Gefahr die Wahrscheinlichkeit eines Unfallereignisses nach vorliegenden Situationsbedingungen zuordnet. Das Unfallereignis ist dabei das Ergebnis der Kombination des initialen Gefahrenereignisses mit den gegenwärtigen situativen Umständen als verkettete Wahrscheinlichkeit.

Dieser soeben erläuterte prospektive Ansatz der Modellierung wird in der *Safety Assessment Methodology* nach Eurocontrol beschrieben (SAM-TF, 2004) und verweist bei der Abstufung der Schweregrade auf die Klassen definiert in *EC No. 1035/2011* (The European Commission, 2011). Die Unterscheidung wird anhand beobachtbarer Indikatoren im Betrieb vorgenommen, die sich z.B. auf die Unterschreitung von Mindeststaffelungswerten beziehen. Die Risikoanalyse umfasst elementar die folgenden Arbeitsschritte:

1. Gefahrenidentifikation
2. Bestimmung der Gefahrenfolgen
3. Bestimmung des Schweregrades

Üblicherweise werden diese Schritte auf Grundlage von Expertenaussagen vorgenommen. Die Schätzungen und Aussagen von Experten zur Bestimmung der Wahrscheinlichkeiten basieren somit i.d.R. auf dem Erlebten und Erfahrenen des Flugbetriebes und stellen somit Schlussfolgerungen auf Basis beobachteter Einzelfälle dar. Die Ereignisbaumanalyse folgt somit der Logik des induktiven Schlusses. Allerdings unterliegen die Expertenaussagen relevanten Störeffekten. Zu diesen gehören z.B.

- subjektive Wahrscheinlichkeitsschätzung (Tversky & Kahneman, 2000), die jedem Menschen zahlreiche Einschränkungen unterstellt, auf Basis des Vergangenen und des Beobachteten valide Wahrscheinlichkeitsschätzungen abgeben zu können. Hierzu gehören z.B. der Effekt des „Bias“ oder eine hohe Unsicherheit, die Stichprobenanzahl und die Grundgesamtheit korrekt einzuschätzen. Die Folge ist eine fehlende Kalibrierung und eine hohe Toleranz der Aussagen (Druzdel & Van Der Gaag, 2000).

- subjektive Bestimmungen der relevanten Faktoren, die für das resultierende Risiko maßgeblich sind. Das Zusammentragen der wesentlichsten Faktoren ist ein zeitraubender Vorgang (Brooker, 2010) und häufig unvollständig.
- Beeinträchtigung der Beobachtung durch Gedächtnisfehler und Falschinformationen, wie sie bei der Untersuchung der Erinnerungen von Augenzeugen nachgewiesen werden konnte (Mudd & Govern, 2004).
- Expertengestützte Risikoanalysen neuartiger Systeme unterliegen einer nicht näher bestimmbaren Beeinflussung durch die mit der Einführung des Systems verbundenen Änderungen der Arbeitsbedingungen. So kann eine signifikante Änderung an z.B. Instrumenten, Technik und Verfahren die Zusammenhänge des Betriebsablaufes neu ordnen, in die sich das Erfahrungswissen nicht übertragen lässt. So könnte ein Experte unter diesem Begleitumstand seinen Status als Träger gültigen Erfahrungswissens verlieren.

Diese Störeffekte führen zu einer begrenzten Validität der Aussagen. Jedoch wird die Verwendung der gewonnenen Informationen in der Praxis häufig gebilligt und durch den Mangel an Alternativen legitimiert:

„Because of the sparsity of relevant empirical data, expert opinion is frequently used in risk analysis to assess the frequencies of rare events” (Mosleh, B., & Apostolakis, 1988, S. 64).

Die Technik hat sich zur Gewinnung von glaubwürdigem aber nicht beweisfähigem Wissen über mögliche Gefahren weiterentwickelt (de Jong, 2004). Für die Risikoanalyse verspricht die Verwendung von Expertenaussagen die effektive Identifikation einer umfänglichen Anzahl an möglichen Gefahrenereignissen.

1.3.2 Modellbasierte Methoden

Die Rekonstruktion des Gesamtbetriebes durch die Modellierung des Einzelverhaltens des menschlichen Operateurs hat das Ziel, auf Basis bekannter Verhaltensmuster quantifizierbare Wahrscheinlichkeiten mittels Simulationstechnik zu gewinnen. Der Vorteil ist die Anwendung von Schnellzeitsimulationstechnik, mit der eine statistisch starke Stichprobenanzahl in einer kurzen Zeit gewonnen werden kann. So kann z.B. Stochastik des menschlichen Verhaltens als Parameter abgebildet werden. Im Rahmen einer Monte-Carlo-Simulation kann sodann die Unsicherheit menschlichen Handelns mit diesem Parameter abgebildet und zu einem daraus resultierenden betrieblichen Risiko quantifiziert werden. Insbesondere der Nachweis sehr kleiner Zielwahrscheinlichkeiten für Un- bzw. Vorfalle in der Luftfahrt kann somit im Rahmen der Modellgenauigkeit verifiziert werden.

Im Bereich der Luftfahrt ist diesbezüglich die Methodik *Traffic Organization and Perturbation AnalyZER* (TOPAZ) repräsentativ, bei der die Handlung des Menschen und ihre Zusammenhänge mit Betriebsereignissen auf Mikroebene auf Basis von Expertenaussagen modelliert werden (Blom H. , et al., 2001). Die Wahrscheinlichkeiten für den Eintritt von Schlüsselereignissen zur Entstehung von Unfällen werden in ein sogenanntes stochastisch-dynamisches Modell überführt. Dieses Modell beschreibt den Zustand einer betrieblichen Situation und die Wahrscheinlichkeit für mögliche Folgezustände auf Basis des Petri-Netz-Formalismus. Die Wahrscheinlichkeiten bilden die Unsicherheiten des Menschen bei

Handlungen wie z.B. Situationsbewusstsein, Überwachung oder Interaktionsvorgängen zwischen Menschen ab.

Die Schnellzeitsimulation liefert die Unfallwahrscheinlichkeit über ein Stichprobenzeitraum von bis zu 10^9 Flugstunden (fl. Hr) und kann das Risiko hinsichtlich kleinster Ursachenbeiträge quantifizieren. Dies stellt gegenüber anderen Simulationstechniken einen erheblichen Vorteil dar, der auf der Wahrscheinlichkeitsachse anhand des *ATM Safety Icebergs* veranschaulicht werden kann (Abbildung 5).

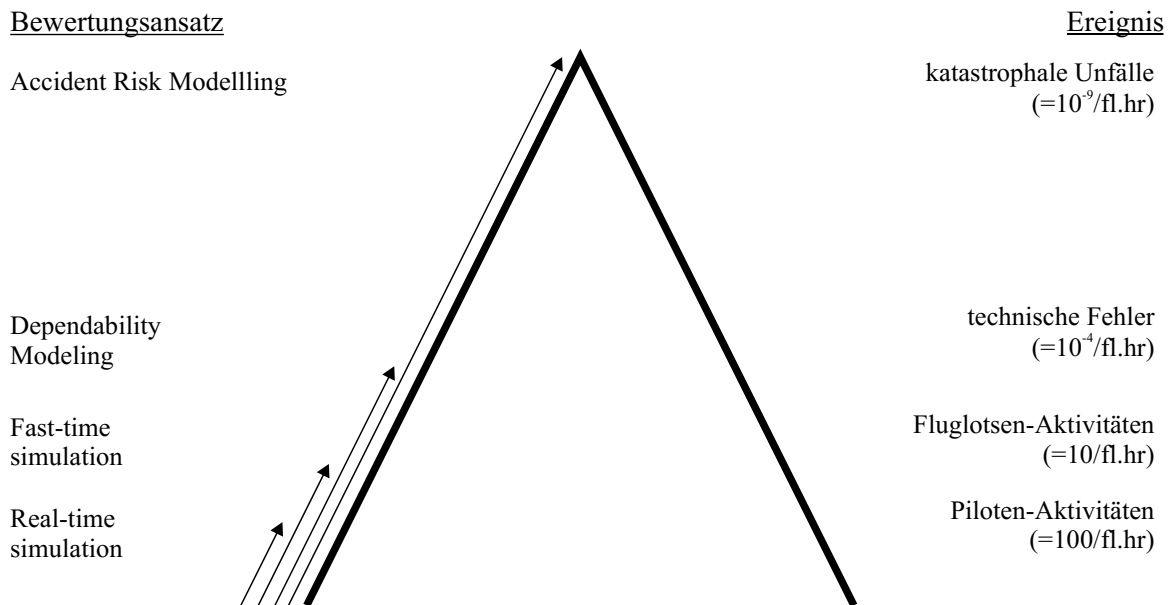


Abbildung 5: ATM Safety Iceberg (Blom H. , et al., 2001)

Der Zusammenhang zwischen der Abweichung vom Anflugpfad bei Anwendung von *Area Navigation* (RNAV)-Verfahren und der Zeitverzögerung durch bordseitige Reaktionszeiten wurde in Bezug auf das entstehende Kollisionsrisiko analysiert (Vogel, Thiel, & Fricke, 2012). Durch die Anwendung von Softwareagenten in der Schnellzeitsimulation konnte das Überschreiten des derzeit als maximal tolerablen Risikoniveaus für fatale Unfälle in Bodennähe von 10^{-7} pro Anflug bei einer Verzögerung zwischen 10 und 20 Sekunden bestimmt werden.

Der Vorteil dieses Ansatzes liegt in der Einbindung komplexer Interaktionsvorgänge zwischen den Operateuren, deren Zusammenspiel eine signifikante Erhöhung des Risikos bewirken kann (Hollnagel, 1993). Ebenso können neuartige Verfahren und Technik unterstellt werden. Die Modellierung setzt dabei A-priori-Wissen über das Elementarverhalten sowie ihre Unsicherheiten voraus und bestimmt somit über die Gültigkeit der erzielbaren Ergebnisse (Callantine, 2003). Insbesondere die Eintrittshäufigkeit von menschlichen Fehlern erfordert nach wie vor die Einbindung von Expertenaussagen. Die Zurückführung des Unfalls auf seine Ursachen ist somit aufgrund des unterstellten Modellwissens hinsichtlich der empirischen Beweisführung umstritten (Hollnagel, 1993).

1.3.3 Die experimentelle Stresstestanalyse

Als besondere methodische Herausforderung erwiesen sich die beispielhaft aufgeführten Ermittlungen zur zunächst ungeklärten Unfallserie der *Comet* von *De Havilland* in den Jahren 1953 bis 1954, deren Überreste nach drei Abstürzen keinen erkennbaren Rückschluss auf die möglichen Ursachen boten (Ministry of Transport and Civil Aviation, 1955). Erst ein experimenteller Versuch bot Aufschluss. Ein intakter Rumpf dieses Modells wurde dabei in einen Wassertank gelegt (Abbildung 6, links). Ein sogenannter zyklischer Belastungstest (eng. *repeated-loading test*) setzte die Druckkabine zyklisch unter Druck. Nach 3.060 Zyklen, deren Belastung auf die Kabine 33 % höher waren als die Normalbelastung unter Bedingungen des Realbetriebes und 2.500 Flügen entsprechen, konnte ein systematischer Konstruktionsdefekt am Rumpf festgestellt werden (Abbildung 6, rechts). Es konnte der Schluss gezogen werden, dass die Lebenszeit der Druckkabine aufgrund höherer Ermüdungseffekte deutlich geringer ausfiel als vom Hersteller geschätzt worden war.

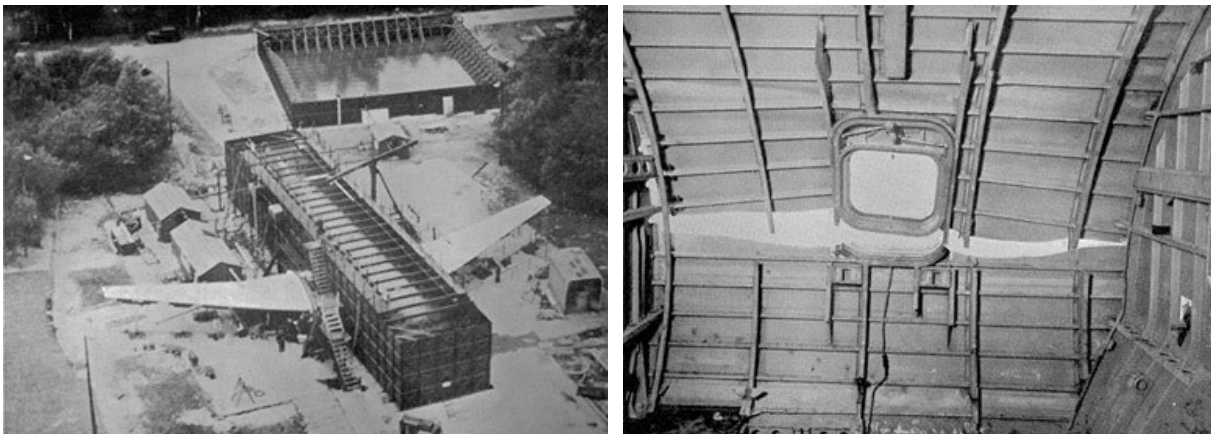


Abbildung 6: Die Comet G-ALYU im Wassertank des Versuchsstandes (links), Rumpfddefekt im Bereich des vorderen Fluchtklappenfensters (rechts) (Ministry of Transport and Civil Aviation, 1955)

Sowohl das Design des modernen Luftfahrzeugs mit Druckkabine, als auch die Methodenentwicklung der Unfallanalyse in der Luftfahrt konnten durch diesen Fall erheblich profitieren: Der Ursachenbeweis *quod erat demonstrandum* unter Laborbedingungen ist das Ergebnis der Anwendung eines deduktiven Untersuchungsansatzes, mit dem die vermutete Unfallursache nach Auftreten des Schadens bestätigt wurde. Beschleunigte Materialermüdungstests (engl. *accelerated fatigue testing*) sind heute ein Standard zur systematischen Identifikation von mechanischen Schwachstellen an Luftfahrzeugen. Als aktuelles Beispiel können die Flügeltests der Industrieanlage Betriebsgesellschaft mbH (IABG) in Kooperation mit der Materialforschung und Anwendungstechnik GmbH (IMA) am Airbus A380 in Dresden genannt werden (Jörn, 2010). 182 Hydraulikzylinder simulierten in den Jahren 2005-2007 die Bedingungen eines Flugzeuglebens von 20–25 Jahren in 26 Monaten.

1.3.4 Die Unfalluntersuchung

Die Ermittlungsarbeiten durch die nationalen Unfalluntersuchungsbehörden sind nach *ICAO Annex 13* (ICAO, 2010) verpflichtend für Luftfahrzeugunfälle und auch bedeutsame Vorfälle.

Somit soll neben den bekannten möglichen Unfallursachen wie z.B. Wetterfaktoren und der Beitrag menschlichen Handelns weitestgehend unbefangen von bislang geltenden Prämissen und Denkparadigmen ermittelt werden. Das Ziel dieser Analyse ist die Prävention von zukünftigen Unfällen durch Ableitung von Sicherheitsmaßnahmen, soweit erkennbar. Die zur Analyse verwendeten Daten sind Aufzeichnungen, die z.B. den Flugdatenschreiber und die Sprachaufzeichnungen im Cockpit umfassen, aber auch Spuren wie z.B. die örtliche Verteilung der Wrackteile. Die Verwendung derartiger Daten sichert die Gültigkeit des Analyseergebnisses als reales Abbild des Flugbetriebs. Nach methodischen Gesichtspunkten erfüllen die auf dieser Basis gewonnenen Daten den wertvollen Anspruch der empirischen Beweisführung für die Bewertung des menschlichen Beitrages zur Unfallentstehung:

„Empirical data are the sine qua non of the Human Reliability Assessment.[...]. The two primary sources of empirical or historical data are: (1) Incident Reporting System, (2) Human Error Rates” (Hollnagel, 1993, S. 122).

Die mögliche kausale Verkettung des Unfallherganges wird dabei anhand des Ausschlussprinzips eingegrenzt, in dem alternative Erklärungen ergründet und ausgeschlossen werden. Auf diesem Prinzip z.B. fußt die Ermittlung des menschlichen Beitrags zum Unfall gemäß der Empfehlungen der ICAO:

„Having established all of the important factual issues making up the occurrence, the investigator must then develop the causal links. All reasonable hypotheses should be stated and evaluated to demonstrate that alternative explanations of the events have been carefully considered” (ICAO, 1993, S. 31).

Die Eingrenzung der wahrscheinlichsten Erklärung wird von der weitreichenden standardisierten Arbeitsweise aller Operateure im Flugbetrieb unterstützt. Die Unfalldatenbanken werden z.B. durch die nationale Untersuchungsstelle gemäß der Richtlinie 94/56/EG gepflegt und setzen somit das obligatorische Meldeprinzip und die Veröffentlichungspflicht für sicherheitsrelevante Ereignisse in der Luftfahrt gemäß *ICAO Annex 13* um (ICAO, 2010). Im Rahmen der Anwendung des *Safety Management Systems* im Flugsicherungsbetrieb (ICAO, 2013b) und auch im Flughafen- und Flugbetrieb sind Datenbanken zu pflegen, die nach dem *ICAO Safety Management Manual* (ICAO, 2013c) einen vorgeschriebenen Umfang und Detailgrad aufzuweisen haben und somit eine Risikoanalyse nach einem vordefinierten Umfang unterstützen können.

Diese Methode der Risikoanalyse bildet für die fortwährende und beständige Evaluation des Gesamtsystems eine effektive Grundlage, das Risikoniveau zu erfassen und geeignete Verordnungen bzw. Sicherheitsbestimmungen zur effektiven Risikominderung zu entwickeln (Abbildung 7):

„Being an important reactive component of the elements contained in the SMS and SSP frameworks, accident investigations contribute to the continuous improvement of the aviation system by providing the root causes of accidents/incidents and lessons learned from analysis of events.” (ICAO, 2013c, S. 34).

Durch diesen retrospektiven Evaluationsmechanismus entsteht als Konsequenz eine Wechselwirkung zwischen der Beobachtung und den daraus ergriffenen Maßnahmen, die es

erlaubt, ändernde Betriebsbedingungen zu berücksichtigen, um den gesetzten Zielsicherheiten zu entsprechen. Der theoretische Effekt dieses reaktiven Kreislaufes wäre ein kontinuierliches Streben des Risikos gegen Null. Der Wissenszuwachs über mögliche zukünftige Unfallentstehungen und die dazugehörigen Risikoschwerpunkte in der Luftfahrtsicherheit ist bislang durch einen damit verbundenen Induktivschluss über den Unfallhergang und die allgemeinen Ursachen leistbar. Diejenigen Unfälle, die trotz der hohen risikomindernden Hürden dennoch auftreten, weisen eine Zusammensetzung von besonderen Faktoren aus:

„They (the accidents) results from a combination of factors, none of which can alone cause an accident, or even a serious incident; therefore, these combinations remain difficult to detect and to recover using traditional safety analysis logic.“ (Amalberti, 2001, S. 112)



Abbildung 7: Der Evaluationskreislauf der Luftfahrtsicherheit

Eine sehr geringe Unfallhäufigkeit in der kommerziellen Luftfahrt, wie typischerweise ein fataler Unfall pro 10 Mio. Flugbewegungen in Europa führt zu hohen Anforderungen an die Identifikation von gemeinsamen Schnittmengen in den Unfallursachen. Zusammenhänge des Unfallherganges und das daraus zu ermittelnde Unfalleintrittsmuster fallen zunehmend generischer aus, wenn eine abnehmende Anzahl von Unfällen berücksichtigt wird. Das daraus resultierende *Induktionsproblem*⁶ macht sich dann an jener Stelle bemerkbar, an der scheinbar Unfallbedingungen auftreten, aber tatsächlich kein Unfall eintritt. Zum einen werden darauf aufbauende Gegenmaßnahmen zwangsläufig zunehmend unschärfer und somit aus Mangel an Bemessungskriterien restriktiver ausgelegt. Zum Zweiten folgt daraus, dass Maßnahmen heutzutage nicht innerhalb einer praktikablen Zeit hinsichtlich ihrer Effektivität verifiziert werden können (Amalberti, 2001). Die Analyse wird zudem durch den stetigen Wandel im Flugbetrieb erschwert, der keine gleichbleibenden bzw. kontrollierbaren Randbedingungen für die Ursachenanalyse gewährt. Ein zunehmender Bemessungszeitraum der zur Analyse hinzugezogenen Unfall- und Vorfallberichte würde somit eine zunehmende Variabilität der betrieblichen Randbedingungen zur Folge haben. Die Effekte dieses „Stichprobenproblems“ werden demnach mit abnehmender Anzahl von untersuchten Unfällen pro Jahr zunehmend stärker.

⁶ Das Induktionsproblem bezeichnet eine Schwäche des Induktionsschlusses, weil der Schluss die Basis des konkret Beobachtbaren und logisch Eindeutigen verlassen (Bortz & Döring, 2006). D.h., es lässt sich nicht formal angeben, wie ein sicherer Schluss auf unbeobachteten oder zukünftigen Ereignissen jenseits von Wahrsagerei möglich sein soll.

Die Effektivität dieses Kreislaufes wird deutlich am Beispiel erfolgreich praktizierter Risikominderung. So konnte z.B. die *Air Line Pilots Association* (ALPA) nachweisen, dass mit 31 aufgetretenen Vorfällen des Typs Runway Incursion der ersten FAA-Kategorie A und B⁷ im Jahr 2006 auf U.S.-amerikanischen Flughäfen die geltenden Sicherheitsziele nicht erfüllt wurden (ALPA, 2007). Die Ursachenanalyse und die Einführung von Sicherheitsmaßnahmen konnten diesem inakzeptablen Risiko erfolgreich entgegenwirken (ICAO, 2007a). Allerdings führen auch hier die Probleme mangelnder Stichprobenzahl zu Schwierigkeiten, die Effektivität der Maßnahme innerhalb einer praktikablen Anzahl an Jahren überprüfen zu können.

Die soeben aufgeführten Probleme mangelnder Stichproben, der die Entwicklung und die Verifikation von risikomindernden Maßnahmen unterliegen, entpuppen sich in der Luftfahrt zunehmend als ein innovationshemmender Faktor. Dies kann mit Hilfe der Zusammenwirkung von Innovation, Sicherheitsmargen und dem Erlass neuer Verordnungen und Sicherheitsbestimmungen erklärt werden. Bei dem Design neuer Verfahren, Trainingskonzepte oder Systeme entsteht gemäß (Amalberti, 2001) eine Lösungsmenge, die allgemein innerhalb von drei Grenzen liegt (Abbildung 8):

- den individuellen und sozialen Verordnungen,
- den Marktregeln und
- den Sicherheitsregeln- und verfahren.

Die Entwicklung eines neuen Systems hat das allgemeine Bestreben, das bestehende System leistungsfähiger und profitabler für das Individuum (individueller Mehrwert) zu gestalten. Obwohl die Risikoanalyse die Erfüllung jeglicher Zielsicherheiten prognostiziert, entsteht durch das neue System ein zusätzliches Risiko, das die bestehenden Sicherheitsmargen zu Gunsten der höheren Leistungsfähigkeit schmälert. Regulierende Instanzen (z.B. nationale Zulassungsbehörden) reagieren mit der Einführung neuer Regeln und Verfahren, um die Sicherheitsmargen aufzubauen bzw. wiederherzustellen. Es entsteht in Konsequenz ein Kreislauf, der überregulierend wirkt und von Amalberti als „Falle der Überregulierung“ bezeichnet wird.

Das Modell nach Amalberti erklärt anhand einfacher Mechanismen, welchen Effekt der Evaluationskreislauf und die damit verbundene Maßnahmenenergreifung hervorrufen können. Er ist für die sogenannten „Ultra-Safe“-Industrien charakteristisch. Falls dieser Kreislauf nicht eine ebenso große Anzahl an Verordnungen hinsichtlich ihrer ursprünglich gedachten Effektivität überprüft und für eine Modifikation oder Aufhebung zur Disposition stellt, entsteht ein Evaluationsprinzip, das additiv wirkt. Aufgrund der damit einhergehenden restriktiveren Auslegung des Flugbetriebs und der Verkleinerung von Spielräumen der Gestaltungsfreiheit macht eine zunehmende Regulierung das Luftfahrtsystem zunehmend träge und wenig anpassungsfähig bezüglich sich ändernder Randbedingungen des Flugbetriebs und der Einführung neuartiger Entwicklungen:

⁷ FAA Runway Incursion Severity Categories. A: Der Abstand verringert sich dramatisch und alle Beteiligten unternehmen extreme Anstrengungen um die Beinah-Kollision zu verhindern oder es folgt ein Unfall. B: Der Abstand verringert sich und es ist ein Potential für einen Unfall vorhanden.

„The rate of production of new guidance materials and rules in the European Joint Aviation Regulations is significantly increasing while the global aviation safety remains for years on a plateau at 10^{-6} (over 200 new policies/guidance/rules per year). Since nobody knows really what rules/materials are really linked to the final safety level, the system is purely additive, and old rules and guidance material are never cleaned up. No surprise, regulations become inapplicable sometimes, and aviation field players exhibit more and more violations in reaction to this increasing legal pressure.“ (Amalberti, 2001, S. 111)

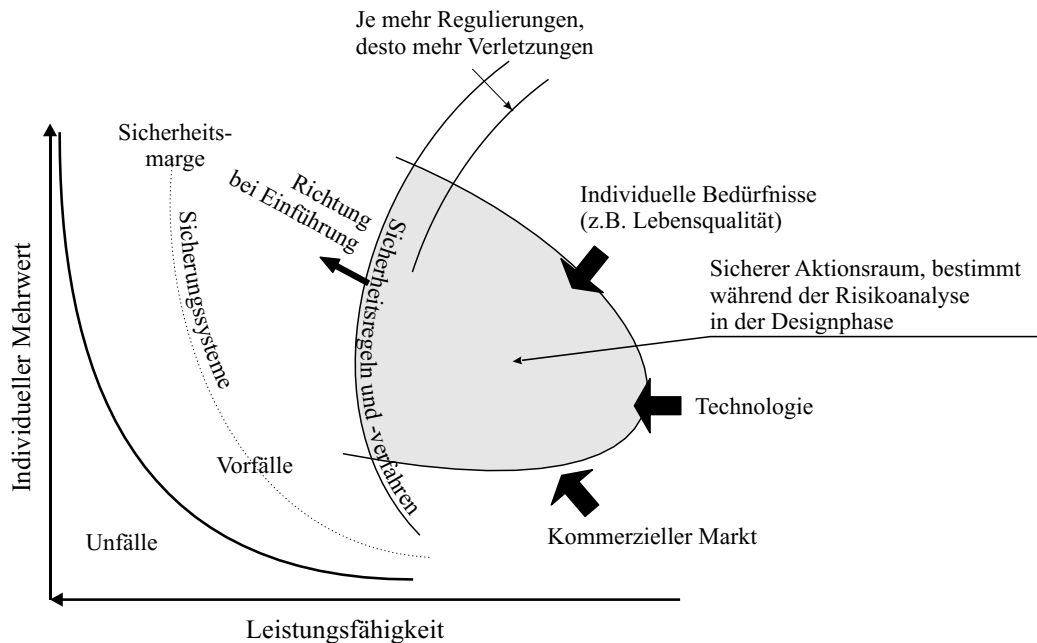


Abbildung 8: Die Falle der Überregulierung (Amalberti, 2001)

Eine sicherheitsorientierte Industrie, welche die Unfallprävention auf Basis dieses reaktionären Prinzips betreibt, bewirkt in Konsequenz eine zunehmende Verlängerung der Entwicklungszeit, um eine Innovation allen vorgeschriebenen Bedingungen anzupassen. Paradoxerweise liegen die Wurzeln dieses Problems in dem bereits erreichten hohen Sicherheit und dem damit verbundenen Stichprobenproblem.

Eine Evaluation der geltenden Verordnungen und Sicherheitsbestimmungen hätte zur Konsequenz, Risiko zulassen zu müssen, um die Effektivität anhand von Betriebsbeobachtungen überprüfen zu können. Eine derartige Feldstudie wäre der Öffentlichkeit aufgrund des experimentellen Charakters und der ethischen Problematik nicht vermittelbar und begründet durch die Unlösbarkeit dieses Problems das Sicherheitsparadoxon aller sicherheitsrelevanten Industrien, die Luftfahrtindustrie eingeschlossen.

⁸ Anmerkung des Autors: Bei dem quantifizierten Wert zum Ausdruck des aktuellen Risikos bezieht sich Amalberti vermutlich auf die relative Häufigkeit von Unfällen zur Menge von Flugbewegungen (als sogenannte Unfallrate) in Europa.

1.3.5 Das „Accident-Incident-Model“

Als Projekt namens *Integrated Risk Picture* (IRP) von Eurocontrol initiiert (Eurocontrol, 2006b) wurden Unfälle wie etwa die *Mid-Air Collision* und *Runway Collision* und Vorfälle mit Hilfe einer Fehlerbaumlogik analysiert. Dies hatte zum Ziel, Risiken mit ATM-Bezug klarer zu verstehen und das Zusammenwirken von Unfallursachen zu demonstrieren. Der Unfallhergang konnte als logische Verkettung mit eindeutigen Ursachenklassen erfasst werden. Möglichkeiten zur Minderung des Risikos können somit unter Berücksichtigung einer allgemeinen Unfalleintrittslogik systematisch identifiziert bzw. optimiert werden. Für die Quantifizierung von Wahrscheinlichkeiten wurden weltweite Un- und Vorfalldaten von 1980 bis 2005 verwendet.

Die Fehlerbaumanalyse IRP wurde in das sogenannte *Accident-Incident-Model* überführt, welches im Rahmen des *SESAR WP 16.1.1.1* im April 2012 als vorläufige Version v0.2 veröffentlicht wurde (SESAR, 2012; SESAR, 2013). Die Fehlerbaummodelle wurden auf Basis von Unfall- und Vorfalldatenbanken der kommerziellen Luftfahrt von acht europäischen Flugsicherungen validiert. Ziel ist es, für die jeweiligen Unfallklassen und seine Ursachen europaweit gültige Zielsicherheiten anhand der Fehlerbäume zu dokumentieren. Zielsicherheiten können anhand solcher Fehlerbäume definiert werden, die zur Verifikation des Risikos bezüglich einer bestimmten Unfallklasse dienen können. Ebenso kann der zeitliche Trend von dokumentierten Ursachenklassen bestimmen werden. Eine effektive Überwachung des aktuellen Risikos ist mit Hilfe dieser Modelle möglich.

1.3.6 Die Problemanalyse

Die gewählten Prämissen sind es, welche für die erreichbare Qualität und somit die Vorhersagestärke des Ergebnisses der Risikoanalyse entscheidend sind. Im Zentrum der Diskussion dieses Abschnitts steht, welche Probleme bei der Risikoanalyse soziotechnischer Systeme den fundamentalen Prämissen der verfügbaren Methoden geschuldet sind. Die Diskussion umfasst hierfür sowohl jene Systeme, die in Betrieb sind, als auch neuartige Systeme. Eine übersichtliche Erklärung für die Probleme liefern zwei Merkmale, nach denen die Methoden grundsätzlich unterschieden werden können.

1.3.6.1 Induktiv vs. Deduktiv

Diese Merkmale beschreiben die Richtung des logischen Schlusses zwischen Allgemeinheit und Einzelfall (Bortz & Döring, 2006). Die Folgerung eines allgemeinen Schlusses nach Beobachtung eines oder mehrerer Ereignisse, z.B. in der Umwelt oder im Flugbetrieb, klassifiziert die Induktion. Die Induktion zeichnet im besonderen Maße aus, Wissen bzw. Vermutungen über potentielle Gefahrenquellen des Unfalls zu schlussfolgern. Die Fehler- und Ereignisbaumanalysen basieren beide auf diesem Prinzip, den allgemeinen Zusammenhang zwischen der Ursache und dem Unfall bzw. anderen Schweregraden nach Beobachtung schlussfolgern zu können. Die Verwendung von Betriebsdaten bietet Vorteile bezüglich der externen Validität. Von Nachteil ist es aber, dass die Verwendung von Betriebsdaten nur jene Systeme zu analysieren gestattet, die bereits in Betrieb sind oder von denen andere Beobachtungen vorliegen.

Die Beweisführung zur Verifikation einer aus der Theorie abgeleiteten Hypothese anhand von Beobachtungen klassifiziert demgegenüber die Deduktion. Einen typischen

Anwendungsfall stellt die Verifikation von Risikomodellen anhand von Betriebsdaten dar. Dies ist der Fall, wenn z.B. sicherheitsrelevante Verschleißteile des Flugzeugs oder Flugdaten anhand von sogenannten *Safety Performance Indicators* überwacht werden und mit einer maximal akzeptablen Eintrittshäufigkeit permanent verglichen werden. Gefährliche Trends können somit frühzeitig erkannt und risikomindernde Maßnahmen rechtzeitig ergriffen werden.

Die Unterscheidung dieses Merkmals liefert eine wichtige Erklärung für die Probleme bei der Identifikation von Gefahren eines neuartigen Systems anhand von Expertenaussagen nach Vorbild der soeben vorgestellten *Safety Assessment Methodology*. Wenn keine Betriebserfahrung mit dem System vorliegt, obliegt es der Vorstellungskraft des Experten, mögliche Gefahren aus dem Allgemeinen abzuleiten. Aufgrund der Erfahrung mit dem konventionellen System führen die vorgeprägte Vorstellungskraft und die Erwartungshaltung des Experten vermutlich zu einer Befangenheit.

1.3.6.2 Retrospektiv vs. Prospektiv

Dieses Merkmal gibt Auskunft über die Untersuchungsrichtung, die von einem Unfallereignis (allg. als Zielereignis bezeichnet) deterministisch rückwärtsgerichtet (retrospektiv) oder zur Bestimmung von potentiellen Folgen von einem Initialereignis (z.B. von einem Gefahrenereignis) probabilistisch vorwärts (prospektiv) gerichtet sein kann. Der Unterschied dieser Merkmale wurde hierzu bereits durch die Modelle in Abbildung 4 illustriert. In der Risikoanalyse ist es folgerichtig ein markantes Unterscheidungsmerkmal retrospektiver Analysen, dass die Ursachen den Eintritt eines Zielereignissen bedingen, wie z.B. ein Unfallereignis.

Das retrospektive Merkmal zeichnet ebenfalls aus, dass es aufgrund fehlender Kontrolle über die Bedingungen zur Erhebung der Daten keine Verifikation einer Kausalhypothese leisten kann. Der Grund ist die fehlende Möglichkeit, Alternativerklärungen ausschließen zu können und ist somit dem prospektiven Untersuchungsmerkmal vorbehalten. Allerdings ist es bei Unfall- und Ursachenanalysen nach *ICAO Annex 13* üblich, Erklärungsmöglichkeiten anhand von Expertenaussagen und anderen Spuren des Unfallhergangs nach Wahrscheinlichkeit und Plausibilität auszuschließen. Die geringe Stichprobenzahl von Unfällen und Vorfällen im Betrieb erschwert die Verifikation von risikomindernden Maßnahmen. Dieses Problem wurde bereits in Abschnitt 1.3.4 unter Verweis auf (Amalberti, 2001) erläutert und kann auf das retrospektive Untersuchungsmerkmal zurückgeführt werden.

Die prospektive Analyse zeichnet demgegenüber aus, dass das Unfallereignis den Eintritt eines ursächlichen Initialereignisses bedingt. Das prospektive Merkmal führt bei neuartigen soziotechnischen Systemen zu dem Problem, das Wissen über mögliche Gefahrenereignisse a priori ermitteln zu müssen.

1.3.7 Die Schlussfolgerung

Der effektiven Unterstützung der Aktivitäten zum Aufbau eines sogenannten Safety Cases⁹ gemäß *Eurocontrol Safety Case Development Manual* (Eurocontrol, 2006a) durch geeignete Methoden stehen die folgenden Probleme bei der Risikoanalyse neuartiger soziotechnischer Systeme entgegen. Die Probleme werden entsprechend der Safety Argumente der jeweiligen Phase des Projekt-Lebenszyklus vorgestellt. Der zu den Argumenten benötigte Sicherheitsnachweis ist bei Änderung der Kontrolldienste der Flugsicherung (z.B. neue Hardware/Software, Verfahren, Training, Instruments, Konzeptänderungen) gemäß Annex 2, Kapitel 3.2 der *EC No. 1035/2011* (The European Commission, 2011) für eine Betriebszulassung zu erbringen.

1. *Konzeptentwicklung*: Die Evaluation des Konzeptes nach Sicherheitskriterien basiert auf Argumenten, die eine vollständige Gefahrenidentifikation und eine präzise Bestimmung der damit verbundenen Unfallwahrscheinlichkeit voraussetzen. Falls das Konzept keine fundamentalen Änderungen umfasst, können für die Gefahrenidentifikation Ereignisbäume wie z.B. das *Accident-Incident-Model* und Unfall- sowie Vorfallberichte verwendet werden. Sie enthalten einen hinreichenden Hinweis über mögliche Risikoschwerpunkte, die es durch geeignete Maßnahmen zu adressieren gilt. Bei fundamentalen Änderungen des Konzepts stützt sich die Gefahrenidentifikation allerdings auf Expertenbefragungen, die aus nun hinreichend erläuterten Gründen keine Vollständigkeit bezüglich der Menge der Gefahren und der Einschätzung der Gefahrenfolgen bietet.
2. *Implementierung*: Die Evaluation des Designs und deren Realisierung beziehen sich auf ein mögliches Verfahrensdesign, Luftraumstruktur, Personaltraining und Instrumente, die den gesetzten Zielsicherheiten entsprechen sollen. Die Verifikation (*Direct Evidence*) erfolgt i.d.R. auf Basis von Expertenaussagen. Es ist allerdings möglich, dass modellbasierte Ansätze der Risikoanalyse Auskunft über die Erfüllung von Sicherheitszielen geben können. Ein bedeutender Nachteil kann allerdings dazu führen, dass das Zusammenspiel zwischen dem Menschen, den Instrumenten und dem Gesamtbetrieb zu neuartigen sicherheitsrelevanten Effekten führen, die in der konzeptionellen Auslegung nicht berücksichtigt werden konnten. Dies ist der Fall, wenn z.B. Instrumente bzw. Bedienelemente des Arbeitsplatzes so ähnlich aussehen, dass eine Verwechslungsgefahr von ihnen ausgeht. Ebenso fehlt hier die Möglichkeit, die Effektivität des gewählten Designs auf Basis von Beobachtungsdaten hinsichtlich der Erfüllung der gesetzten Sicherheitsziele zu verifizieren.
3. *Migration*: Neuartige Gefahren, die bei der Transition von dem Vorzustand zu dem neuen Zustand resultieren können, können gegenwärtig ebenfalls nur mit Hilfe von Expertenbefragungen vorhergesagt werden. Die Effektivität möglicher, ggf. entwickelbarer, risikomindernden Maßnahmen kann ebenso nur mit diesem Mittel verifiziert werden.

⁹ Der Safety Case ist der dokumentierte Nachweis (z.B. argumentativ oder mit Beweisführung) über das Erreichen oder die Gewährleistung von Sicherheit. Es handelt sich primär um ein Mittel für jene, die Dienste bereitstellen oder Projekte umsetzen, um sich selbst beweisen zu können, dass der Dienst oder das Projekt unter Maßgabe einer akzeptierten Zielsicherheit bereitgestellt bzw. umgesetzt ist und wird (Eurocontrol, 2006a).

Die soeben aufgeführten Argumente lassen erkennen, dass die Probleme der Risikoanalyse neuartiger soziotechnischer Systeme nicht auf konzeptionelle Schwächen der Methoden, sondern auf den schlichten Mangel an unabhängigen Betriebsbeobachtungen zurückzuführen sind. Dies betrifft in erster Linie die induktiven Methoden wie z.B. die Unfalluntersuchung, denen keine Beobachtungsdaten zugrunde liegen. Aber ebenso sind deduktiven Methoden aufgrund der fehlenden Möglichkeit zur Verifikation betroffen.

1.4 Die Echtzeitsimulation als mögliche Lösung

1.4.1 Der gegenwärtige Nutzen von Echtzeitsimulatoren im Flugbetrieb

Die Entwicklung der Technik und von Modellen zur Abbildung der physikalischen Randbedingungen sind die maßgebenden Taktgeber für die Realitätstreue von Simulatoren in der Luftfahrt. Im professionellen Anwendungsbereich dienen Simulatoren Trainings- und Forschungszwecken. Typischerweise umfassen die Simulatoren die Arbeitsplätze zur Führung des Luftfahrzeugs oder die Fluglotsenarbeitsstation zur Kontrolle des Verkehrs durch die Flugsicherung. Die kontrollierbaren Bedingungen der Echtzeitsimulation entsprechen grundsätzlich einem Laborexperiment, in welchem der Versuchsperson eine künstliche Arbeitsumgebung geschaffen wird. Der Betrieb von Simulatoren weist einen gegenüber dem echten Betrieb sehr wertvollen Vorteil auf: eine risikofreie Arbeitsumgebung.

Der erste Flugsimulator wurde 1917 erwähnt und hatte eine in drei Achsen drehbare Pilotenkanzel mit dem Namen *Ruggles Orientator* (Jones, 1927). Die Lage konnten über das Steuerhorn des Piloten oder des Ausbilders vorgegeben werden (Abbildung 9). Der Ausbilder konnte eine beliebige Lage herstellen, aus der der Pilot in den horizontalen Geradeausflug zurückfinden sollte. Den Anlass hierzu bot zu dieser Zeit eine Serie von Unfällen, die auf die Orientierungslosigkeit von Piloten bei Extremmanövern zurückzuführen war. Für ein vorbereitendes Training derartiger Situationen konnte der Ausbilder die Ausgangslage der Kanzel vorgeben und störend auf die stabilisierenden Bemühungen des Piloten wirken. Der Trainingswert für seltene und gefährliche Flugzustände wurde erkannt. Dieses Training verhalf dem Piloten zu einer verbesserten Kontrolle des Luftfahrzeugs.

Die beständige Weiterentwicklung befähigt die Simulatoren heutzutage z.B. Bewegungen eines Flugzeugs auf einer Bewegungsplattform unter Nutzung sowohl der 3 Rotationsfreiheitsgrade als auch der 3 Translationsfreiheitsgrade als sogenannter *Full Flight Simulator* mit einer hohen gefühlten Realitätstreue nachzubilden (EASA, 2012). Das Grundprinzip, die äußeren Anfangs- und Randbedingungen des Betriebes zu kontrollieren, ist stets gleich geblieben. Sekundäre Störeffekte können systematisch randomisiert oder eliminiert werden. Diese Kontrolle gelingt mit Hilfe einer künstlichen Umwelt, die eine Arbeitsumgebung für die Versuchsperson mit Hilfe

1. von Modellen der physikalischen Umgebungsbedingungen und der Bord- bzw. Bodensysteme, die das Verhalten der Arbeitsumgebung abzubilden im Stande sind, und
2. unter Verwendung gleichwertiger Kommunikationspartner in einer „Multiakteurumgebung“

herstellt. Als Kommunikationspartner (im Fachjargon auch „Gamer“ genannt) können instruierte Operateure wie z.B. Pseudopiloten angeführt werden. Die Instruktionen sind dann Bestandteil des Aufbaus der Arbeitsumgebung und somit Teil eines künstlichen Verhaltens

innerhalb des Experimentalaufbaus. Die Modelle und die Kommunikationspartner binden den Menschen über Bedien- und Anzeigeeinstrumente als handelnde Entscheidungsinstanz in den Kreislauf ein. Der englische Begriff *Human-In-The-Loop-Simulation* trifft das Prinzip dieses Wirkungskreislaufes sehr gut. Der äquivalente deutsche Begriff *Echtzeitsimulation* weist terminologisch auf die Möglichkeit hin, die Modelle ebenso ohne Einbindung des Menschen für eine Schnellzeitsimulation anwenden zu können.

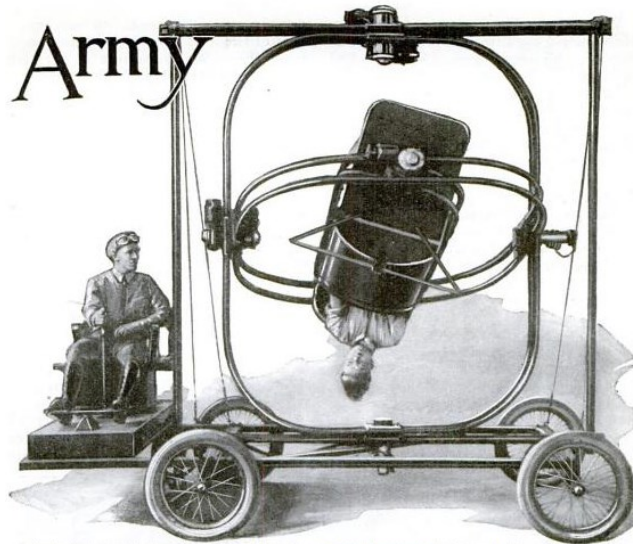


Abbildung 9: Der *Ruggles Orientator* (Windsor, 1928)

Die Entwicklung von Simulatoren in der Luftfahrt ist bislang vornehmlich getrieben von Entwicklungen im Bereich der Computertechnologien, die eine „fotorealistische Realitätstreue“ herstellen (Dahlstrom, Dekker, Winsen, & Nyce, 2009). Ebenso ist die Entwicklung der Bewegungsplattform ein entscheidender Bestandteil für das realitätsgetreue Empfinden des Operateurs. Äquivalent dazu erfährt der Flugbetrieb ebenso eine kontinuierlich zunehmende Automatisierung sowie Computerisierung und ist dem Entwicklungstrend der Simulatoren entgegengerichtet. Der reale und der simulierte Flugbetrieb werden sich demnach zunehmend ähnlicher.

Die technisch leistbare Realitätstreue und die Zertifizierung von Flugsimulatoren anhand von Spezifikationen, wie z.B. *CS-FSTD* gemäß EASA (EASA, 2012), schaffen eine weitreichende Akzeptanz im Bereich des operativen Flugtrainings. Die Zulassungen für die Führung von Luftfahrzeugen und Lotsentätigkeit kann aus diesem Grund mit Hilfe von simulierten Zulassungstests erteilt werden, wie z.B. der *Pilot Proficiency Check* gemäß *ICAO Manual Doc 9994 Manual of Evidence-Based Training* (ICAO, 2013).

Die Kontrolle über die Arbeitsumgebung bietet hier für das Training die wertvolle Möglichkeit, eine Reaktion der zu trainierenden Person mit Hilfe von Gefahrensituationen zu stimulieren. Dies kann z.B. ein Triebwerksausfall sein, bei dem der Pilot die verfahrenskonforme Abarbeitung von Arbeitsschritten zum Ausschalten oder zum Neustart des Triebwerks zeigen soll (Fischer, Orasanu, & Montalvo, 1993). Ebenso bietet es die Reproduzierbarkeit von Gefahrensituationen durch die Herstellung gleicher Anfangsbedingungen für z.B. die Analyse der Qualität des manuellen Anflugs anhand der

Abweichung vom Sollpfad (Haslbeck, Kirchner, Schubert, & Bengler, 2014). So konnte anhand der Abweichung auf einen Zusammenhang zwischen der Qualität zur manuellen Führung des Luftfahrzeugs und der Anzahl praktizierter und trainierter Flugstunden zurückgeschlossen werden. Im Rahmen der Risikoanalyse kann z.B. die Risikosensitivität¹⁰ des Flugbetriebes bei Eintreten einer vordefinierten Gefahrensituation bestimmt werden (Stroeve, Blom, & Bakker, 2013). Der Effekt der Gefahr kann dann unter Kontrolle oder Eliminierung anderer Störvariablen anhand der Primärvarianz¹¹ beobachtet werden.

Die Anwendung der Echtzeitsimulation im Rahmen von Risikoanalysen umfasst sowohl die Identifikation von Ursachenereignisse fehlerhafter menschlicher Handlungen als auch deren Eintrittshäufigkeit (Kirwan, 2007). Simulatoren qualifizieren sich aufgrund ihres risikounbelasteten Funktionsprinzips für die gefahrenlose Erprobung in einer dem Betrieb äquivalenten Umgebung, die es auch erlaubt, neuartige soziotechnische Systeme und Verfahren zu erproben. Aus diesem Grund werden Simulationsstudien für die Evaluation des Systemdesigns verwendet. Sie erlauben einen Rückschluss auf die resultierende Effizienz und Sicherheit des Betriebs auf empirischer Basis (Prevot, Lee, & Callantine, 2010).

Die Einbindung von Simulationstechnik in die Risikoanalyse ist für die Unterstützung von Modellierungsprozessen sinnvoll, in dem einzelne Wahrscheinlichkeiten und Prozessgrößen für die Kalibrierung von Agenten bestimmt werden, die für die anschließende Schnellzeitsimulation zur Anwendung kommen können. Einzelne Konfiguration können innerhalb von wenigen Tagen simuliert werden und können eine Stichprobenanzahl von einigen 100 Stunden liefern (Stroeve, Blom, & Bakker, 2013).

Ein weiterer Vorteil der Echtzeitsimulation ist die Messbarkeit von Gefahrenindikatoren. Diese bieten einen Rückschluss auf mögliche Verkettungen von Unfallursachen, deren Effekt auf das Risiko modellbasiert erklärt werden kann. Hierzu gehört etwa die Messung der visuellen Wahrnehmung mittels Eye-Tracking-Technologie, die die Blickbewegungsmuster von Menschen aufzeichnen können. Diese Technik erlaubt auch für die Risikoanalyse interessante Ansätze wie z.B. Gefahren der verfehlten visuellen Wahrnehmung mit einem Risiko zu verbinden (Koch, 2014). Zu diesen Indikatoren gehören ebenfalls Messungen der Arbeitsbeanspruchung, die durch eine Arbeitsbelastung entsteht und im Überlastfall kritische Situationen provozieren kann. Als Beispiel können untersuchte Komplexitätsfaktoren der Streckenkontrolle aufgeführt werden, bei denen zwischen der entstehenden Arbeitsbeanspruchung und dem Kollisionsrisiko ein Zusammenhang besteht (Djokic, Lorenz, & Fricke, 2008).

1.4.2 Ein Konzept zur simulationsgestützten Risikoanalyse

1.4.2.1 Der Wirkungskreislauf

Wie bereits erläutert stellt die Echtzeitsimulation ein Analysemittel dar, das es ermöglicht, neuartige soziotechnische Systeme schon während des Konzept- und Designprozesses sowie der Erprobungsphase unter kontrollierten und reproduzierbaren Bedingungen zu testen. Der

¹⁰ Risikosensitivität ist die Sensitivität bzw. die Empfindlichkeit des Risikos gegenüber der Änderung eines einzelnen Parameters und Einflussfaktors.

¹¹ Die Primärvarianz bezeichnet diejenigen Veränderungen der Messwerte, die "allein auf die Variation der experimentellen (Instruktions-)Bedingung zurückzuführen" sind (Sarris, 1990, S. 167).

Wirkungskreislauf der Echtzeitsimulation stellt darin den menschlichen Operateur als Entscheidungsträger mit gegebener Aufgabendefinition in den Mittelpunkt (Abbildung 10). Der Wirkungskreislauf, in dem die Handlungen des Operateurs über die Schnittstellen einem virtuellen Arbeitsumfeld zugeführt und die Reaktion des Umfeldes wiederum zum Operateur zurückgeführt werden, wird mit Hilfe von Mensch-Maschine-Schnittstellen geschlossen. Die Simulation wendet dazu Modelle an, die Physik, Technik und Organisation abbilden können, und alle Ereignisse, die Verkehrsbewegungen inbegriffen, über die kontinuierliche Zeitachse beschreiben. Ebenso wurde in der Einleitung die Möglichkeit zur Nutzung von instruierten menschlichen Operateuren erläutert. Die Interaktion mit der virtuellen Arbeitsumgebung schafft die Voraussetzung, den menschlichen Operateur als handelnde Instanz in den simulierten Betrieb einzubinden.

1.4.2.2 Die Datenerhebung

Die Echtzeitsimulation bietet für die Bestimmung des Risikos je nach Anwendungszweck Zugangspunkte für die Erfassung geeigneter Aufzeichnungsgrößen. Es kann nach drei Arten von Größen unterschieden werden (vgl. Abbildung 10):

1. sicherheitsrelevante Ereignisse,
2. Ursachenereignisse und
3. objektive Sicherheitsindikatoren.

Sie werden im Folgenden näher vorgestellt.

Die sicherheitsrelevanten Ereignisse

Die oberste Ebene bildet für die Risikoanalyse die Beobachtung des Unfallereignisses oder auch der zugehörigen Vorfalleereignisse, die unter dem Begriff der sicherheitsrelevanten Ereignisse zusammengefasst werden. Die Anzahl der beobachteten Ereignisse führt in Verbindung mit dem Beobachtungszeitraum der Simulation zu einem bestimmbar Risiko. Diese stellen eine Datenlage zur direkten Bestimmung des Risikos her und sind Ausgangspunkt für eine mögliche Beweissicherheit über das Erreichen einer Zielsicherheit.

Die Ursachen

Die zweite Ebene bilden objektive Messungen von Zuständen oder Ereignissen, deren Gefahrenpotential bekannt ist und ein Bindeglied in der Verkettung des Unfallherganges bilden kann. Hierzu gehören Beobachtungsdaten, die das menschliche Verhalten und die Leistungsfähigkeit erfassen und die einen Rückschluss auf den Beitrag des menschlichen Handelns zur Entstehung der Gefahrensituationen bieten. Als typisches Beispiel können Verfahrensverletzungen genannt werden, wie z.B. das Auslassen von Arbeitsschritten auf der Checkliste. Ebenso können Handlungen zur Wiederherstellung eines verfahrenskonformen Betriebsablaufes erfasst werden. Dies setzt allerdings auch voraus, dass eine kausale Beziehung zur Unfallwahrscheinlichkeit aufgrund des Analyseergebnisses einer Unfalluntersuchung bereits bekannt ist, die einen direkten Schluss auf das Risiko zulässt. Diese können z.B. Reaktionszeiten oder Detektions- bzw. Erkennungswahrscheinlichkeiten des Menschen während des Auftretens einer zeitkritischen Situation sein.

Ebenso können subjektive Einschätzungen des Operators über das Zusammenwirken von situativen Faktoren zusätzliche Hinweise auf die Ursachen bieten. Typische Methoden sind die Nachbefragungen und auch die Anwendung von Tests, um die Leistungsfähigkeit des Operators und die damit verbundene Qualität in einen Zusammenhang mit dem Betriebsablauf und dem gemessenen Risiko setzen zu können.

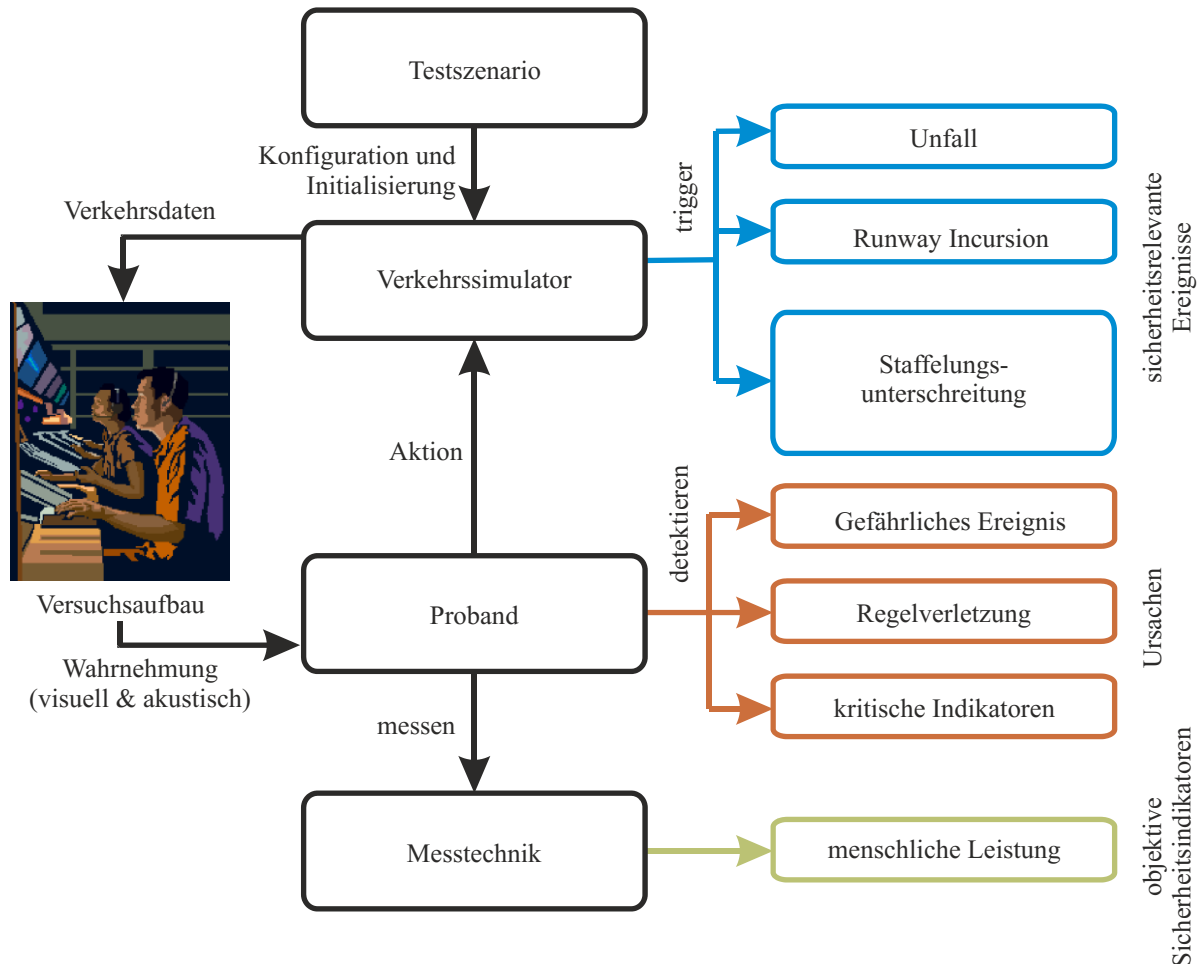


Abbildung 10: Prinzip der Echtzeitsimulation für die Risikoanalyse

Die Sicherheitsindikatoren

Zu der dritten Ebene gehören objektive Messverfahren wie z.B. das *Eye-Tracking*-Messverfahren, mit dessen Hilfe Blickbewegungsmuster des Operators erfasst werden können und die Rückschluss auf einen Zusammenhang mit der Aufmerksamkeit und einer erhöhten oder auch verminderten Qualität der Entscheidungsfindung bieten können. Ebenso können hier beispielsweise physiologische Messverfahren zur Anwendung kommen, um z.B. Herzpulsraten, Stresshormone oder andere Indikatoren zu erfassen. Vornehmlich sollen die Zusammenhänge dabei Aufschlüsse bieten, welche Effekte des Menschverhaltens für das Entstehen von Schaden eine Rolle gespielt haben könnten. Eine kausale Relation zu den übrigen Datenebenen ist dabei nicht zwangsläufig herstellbar, wohl aber die Bestimmung von Grundtendenzen bezüglich des Auftretens von Fehlverhalten. Als konkretes Beispiel kann das visuelle Blickbewegungsverhalten des Platzverkehrsleiters unmittelbar vor der Startfreigabe

genannt werden, das bestimmte Bereiche der Start- und Landebahn abtastet und somit die Pistenfreiheit für einen Startvorgang sicherstellt (Koch, 2014).

1.4.2.3 Das Anwendungsbeispiel „Single Remote Tower“

Das Konzept zur Kontrolle von Luftfahrzeugen auf dem Flughafen aus der Ferne ist bekannt unter dem Namen Remote Tower oder auch Virtual Tower, wobei sich in den letzten 4 Jahren der erstere Begriff durchgesetzt hat. Das Konzept beruht auf der Substitution der Außensicht des Tower-Arbeitsplatzes des Platzverkehrslotsen durch einen Außensichtersatz. Durch die gewonnene Ortsunabhängigkeit ist es möglich, Flugplätze von einer zentralen Leitstelle aus zu kontrollieren. Die Zuordnung eines Flugplatzes zu nur einem Platzverkehrslotsen wird auch unter dem Begriff *Single Remote Tower* geführt. Gegenwärtiges Beispiel ist die Einrichtung eines zentralen Remote Towers in Leipzig (DFS, 2014). Die erste Betriebszulassung für den Betrieb des Remote Towers wurde dem schwedischen Flugsicherungsunternehmen LFV erteilt, das ab dem ersten Quartal 2015 den Flughafen Örnköldsvik aus der Ferne kontrollieren durfte (Nolte, 2014).

Die Arbeit des Platzverkehrslotsen mit Hilfe des Remote Towers könnte durch die Änderung eines wichtigen Arbeitsmittels wie die Außensicht beeinträchtigt werden, da zahlreiche negative Effekte des Außensichtersatzes auf die Qualität der Flugplatzkontrolle noch nicht bekannt oder nicht quantifiziert werden können (Meyer, Vogel, & Fricke, 2010). Vor diesem Hintergrund hat die *European Cockpit Association* (ECA) ein Positionspapier zu dem Konzept veröffentlicht, das Argumente und Sorgen bezüglich der Qualität der zu erbringenden Kontrolle beschreibt (European Cockpit Association, 2014), wie z.B. die Sicherheit vor Hackern, die Qualität der Fluginformationsdienste sowie die ordentliche Feststellung der Pistenfreiheit.

Im Rahmen einer simulationsgestützten Risikoanalyse mit 13 studentischen Probanden wurden mit Hilfe einer Echtzeitsimulation mehrere Außensichtgestaltungen entwickelt (Tabelle 1) und hinsichtlich ihres Beitrages zum Risiko verglichen (Meyer, Schultz, & Fricke, 2011). Hierzu gehörte die Anordnung von zwei möglichen Perspektiven des Außensichtersatzes (Abbildung 11) und die Verfügbarkeit der Bodenlagedarstellung. Die gekreuzte Perspektive bietet eine deutlichere und größere Darstellung der Auf- und Abrollwegen zur Piste und soll die Überwachung dieses kritischen Übergangsbereiches zwischen Piste und Rollwegen durch den Platzverkehrslotsen verbessern. Das unerlaubte Befahren der Piste soll somit schneller und sicherer erkannt werden und der Auftritt von Runway Incursion frühzeitig vermieden werden. Demgegenüber steht die Verfügbarkeit der Bodenlagedarstellung, die anhand einer schematischen Abbildung des Flugplatzes und den aktuellen Positionen und Richtungen der Flugzeuge und Bodenfahrzeuge eine klare Übersicht für den Lotsen bietet. Die Verfügbarkeit soll den Lotsen befähigen, die Wirbelschleppenstaffelung von Luftfahrzeugen vor und hinter der Piste sicherzustellen.

Tabelle 1: Die Konfigurationen als Kombination der variierten Gestaltungsvarianten

	<i>mit Bodenlage</i>	<i>ohne Bodenlage</i>
<i>zentrale Perspektive</i>	A	B
<i>gekreuzte Perspektive</i>	C	

Die Auswertung lässt auf eine intensive Nutzung der Bodenlagedarstellung schließen, die eine starke Minderung der Risiken bei Benutzung der Bahn bewirkt. Die Minderung ist anhand der kontrastreichen und übersichtlichen Vermittlung der Betriebssituation auf dem Flughafen durch diese Anzeige erklärbar. Die daraus resultierende Aufmerksamkeitsverteilung fällt anscheinend zu Lasten der Erkennbarkeit von Ereignissen. Diese sind mittels der Bodenlagedarstellung, aus Gründen der Erfassbarkeit (Tiererscheinungen) oder auch der Perspektive (Durchstartvorgänge) schlechter oder gar nicht erkennbar. Dies lässt sich anhand des Effektes der Head-Down-Zeit erklären, der eine weniger intensive Beobachtung des tatsächlichen Geschehens auf den Betriebsflächen zur Folge hat. Es ist aber auch ein novizenbedingter Effekt zu vermuten, bei dem der permanenten Beobachtung der Außensicht zu wenig Bedeutung beigemessen wurde, die somit keine Integration in den regulären Arbeitsfluss erfährt. So würden vermutlich intensiveres Qualitätstraining und längere Trainingszeiten zu einem Ausgleich der Aufmerksamkeit führen und die Außensicht somit stärker gewichten.

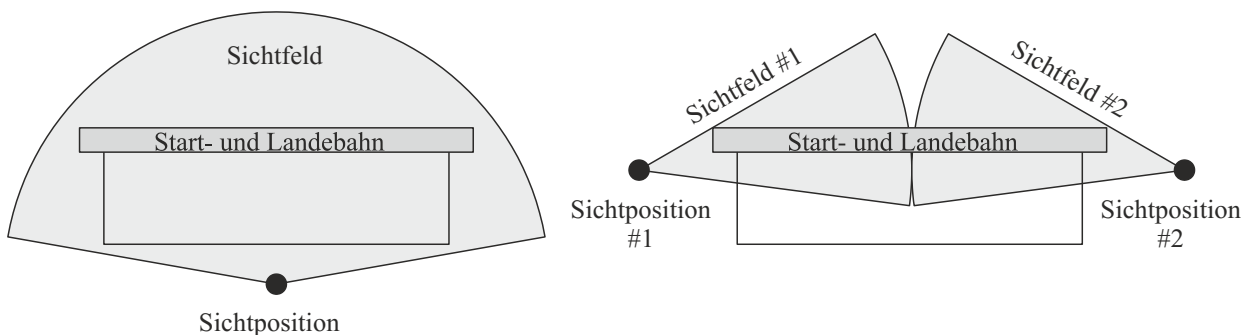


Abbildung 11: Zentralperspektive (links) und gekreuzte Perspektive (rechts)

Die Ergebnisse der gekreuzten Perspektive konnten den erwarteten Mehrwert bei der Erkennung von unerlaubtem Befahren der Piste nicht anhand der Anzahl beobachteter *Runway Incursion* bestätigen (Abbildung 12). Die wenig intuitive Perspektive erfordert vermutlich eine höhere Konzentration durch den Probanden, die einzelnen Perspektiven zu einem mentalen Gesamtbild zu verknüpfen. Die zentrale Perspektive bietet im Gegensatz dazu den Vorteil der Egozentrik, die die Situation auf dem Flugplatz intuitiver vermittelt. Die Bodenlage kann anhand der Anzahl aufgezeichneter *Runway Incursion* und Verletzungen der Wirbelschleppenstaffelung einen enormen risikomindernden Beitrag verzeichnen.

So ist insbesondere die relative Verteilung des mit den Vorfalleignissen verbundenen Risikos deutlich erkennbar. Die Konfiguration A wurde somit als jene Konfiguration identifiziert, die gegenüber den anderen Konfigurationen das kleinste Potential für Eintritt von *Runway Incursion* bietet. Ursachen im Design können anhand einer derartigen Risikoanalyse relativ zueinander gewichtet werden und für die Formulierung von Sicherheitsanforderungen dienen.

1.4.2.4 Das Anwendungsbeispiel „Multi Remote Tower“

Der Remote Tower ebnet aufgrund der Ortsunabhängigkeit dem neuartigen Konzept den Weg, mehrere Flugplätze durch einen Platzverkehrslotsen zu kontrollieren. Dieses Konzept namens *Multi Remote Tower* soll verkehrsarme Flugplätze zu einer Kontrollleitstelle zusammenfassen

und Kontrolldienstleistungen auf Bedarf anbieten. Zu den Vorteilen gehört die verbesserte Auslastung des Platzverkehrslotsen.

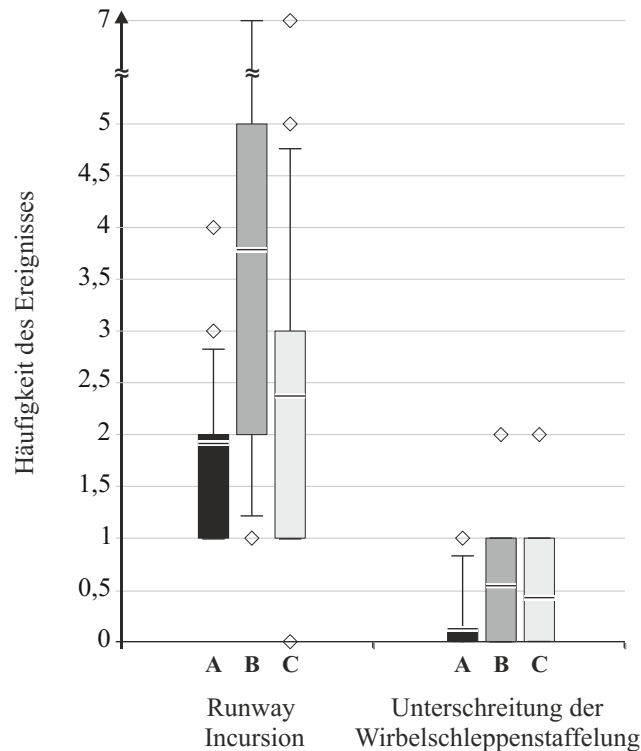


Abbildung 12: Häufigkeiten einer Runway Incursion und einer Unterschreitung der minimal-zulässigen Wirbelschleppenstaffelung

Die konventionellen und bewährten Arbeitsmuster des Platzverkehrslotsen erfahren vermutlich durch ein derartiges Arbeitsschema eine deutliche Änderung. Dieses lässt sich mit dem Arbeiten in einem wechselnden oder auch alternierenden Arbeitsumfeld erklären, bei der jeder Flugplatz eine unabhängige Betriebssituation aufweisen kann. Durch die alternierende Umgebung ist der Platzverkehrslotse gezwungen, wahrgenommenen Informationen über aktuelle und geplante Verkehrsbewegungen sowie bereits erteilte Freigaben einem spezifischen Flugplatz zu zuordnen. Die abgeforderte Gedächtnisleistung ist vermutlich höher als bei dem einfachen Remote Tower, wenn z.B. Windrichtung, -stärke, Wolkenhöhe, Niederschlag, Druck und Flugplatz-spezifische Betriebsabsprachen, als Flugplatz-vielfache Information zu merken und korrekt zuzuordnen sind. Die Namenskonventionen von ICAO-konformen Flugplätzen, deren Rollwegnamen und Meldepunkte einem gleichen oder zumindest ähnlichen Schema folgen, sind dabei der Bereitstellung von Unterscheidungsmerkmalen entgegengerichtet und somit dem Assoziieren von Gedächtnisinformationen vermutlich abträglich. Da Gedächtnisfehler bekannte und gut erforschte Unfallursachen darstellen, ist durch diesen neuen Freiheitsgrad der Information von einer möglichen Beeinträchtigung der Sicherheit auszugehen. Zu diesem Zweck sollen Risiken, die durch den Effekt mangelnder Merkmalsrepräsentation für den Betrieb entstehen könnten, mit Hilfe einer ersten experimentellen Pilotstudie identifiziert werden.

Für die Untersuchung dieses Effekts wurde ein Versuch durchgeführt, der einen Multi Remote Tower-Arbeitsplatz (Abbildung 13) mit drei Pseudopilotenarbeitsplätzen umfasste. Das Versuchsszenario bestand aus drei VFR-Flugbewegungen mit Start-, Lande- und Überflugvorgängen auf den Flughäfen Dresden, Münster-Osnabrück und Dortmund, zwischen denen der Platzverkehrslotse die Sicht wechseln konnte. Es wurden 5 Versuche mit einem Szenario durchgeführt, das je ca. 1 Stunde und 20 Minuten umfasste.

Als Versuchsprobanden wurden 5 Platzverkehrslotsen der *Deutschen Flugsicherung GmbH* und als Pseudopiloten Personen mit einer gültigen *PPL* eingesetzt. Ziel des Versuches war die Beobachtung von Indikatoren, die auf das Verwechseln oder Vergessen von sicherheitsrelevanten Betriebsinformationen hindeuten. Der zu untersuchende Effekt sollte mit Hilfe der Stressreaktion bei Eintritt von Gefahrensituationen stimuliert werden. Der mit der Gefahr verbundene Reiz soll durch den Probanden gemerkt und später dem korrekten Flugplatz zugeordnet werden, um die korrekte Assoziation nachzuvollziehen. Die Simulation konnte hierzu eine Auswahl von Gefahren in den Betrieb einblenden (Tabelle 2), die ordnungsgemäß zu erkennen waren. Der Proband war angewiesen, ungewöhnliche Ereignisse zu erkennen und zu melden. Eine Erkennungsrate der eingeblendeten Gefahren konnte somit ermittelt werden. Ebenso wurden freie Nachbefragungen durchgeführt, die subjektive Aussagen zu möglichen Gefahren protokollieren sollten.

Tabelle 2: Gefahrenereignisse

	<i>Beschreibung</i>	<i>Schweregrad bei Nicht-Wahrnehmung</i>
1.	Tiererscheinung	Unfall
2.	unbefugtes Befahren der Piste durch ein Bodenfahrzeug	Unfall
3.	Unwettererscheinung	Unfall
4.	unangemeldetes Einfliegen in Kontrollzone (Intruder)	erhebliche Störung

Das unbefugte Befahren der Piste ebenso wie die Unwettererscheinung konnte in je einem von 5 Fällen nicht erkannt werden. Der Intruder wurde in 4 von 5 Fällen nicht erkannt. Die Nachbefragung ergab, dass die Probanden die erkannten Gefahren dem Flugplatz in 3 von 20 (15%) der Fälle im Nachhinein nicht korrekt zuordnen konnten. Es gaben weiterhin 2 von 5 zu Protokoll, dass es eine subjektiv hohe Wahrscheinlichkeit für das Verwechseln von ATIS-Informationen gibt. Die Gefahr für die Verwechslung von Rollwegbezeichnungen, Platzrundenrichtung und Luftfahrzeug-Positionen wurden durch je 1 von 5 bestätigt. Dies betrifft insbesondere die Rollwege von und zur Piste, die gleiche Designatoren wie z.B. „alpha“ und „beta“ haben. Die Video- und Audioanalyse ergab je eine Sicht- und Wegpunktverwechslung aber keinmal die Verwechslung des Towerrufzeichens.

Die Ergebnisse und die Aussagen der Probanden weisen anhand einer kleinen Anzahl von Versuchen nach, dass die Gefahr für Verwechslung real und messbar ist. Die Studie verdeutlicht somit, dass das Risiko der Verwechslung existent ist. Allerdings machen derartige Experimentalstudien mit geringen Probandenzahlen und kleinen Zeiträumen allgemein nur einen kleinen Ausschnitt der möglich auftretenden sicherheitsrelevanten Ereignisse sichtbar. Es ist wahrscheinlich, dass eine längere Studie hier zu weitaus

detaillierteren Aussagen befähigen könnte, die z.B. die Risikoschwerpunkte innerhalb der möglichen Informationselemente identifiziert. Es ist ebenso wahrscheinlich, dass der Effekt für wenige Wechsel zwischen den Flugplätzen am Tag gering ausfällt. Als sehr relevante risikomindernde Maßnahme kann die Unterstützung des Arbeitsplatzes durch zusätzliche Unterscheidungsmerkmale erachtet werden. Hierzu können Farb- und Formassoziationen sowie die Modifikation der Namenskonventionen gehören, die dem Platzverkehrslotsen ein charakteristisches Merkmal zur korrekten Zuordnung bieten.

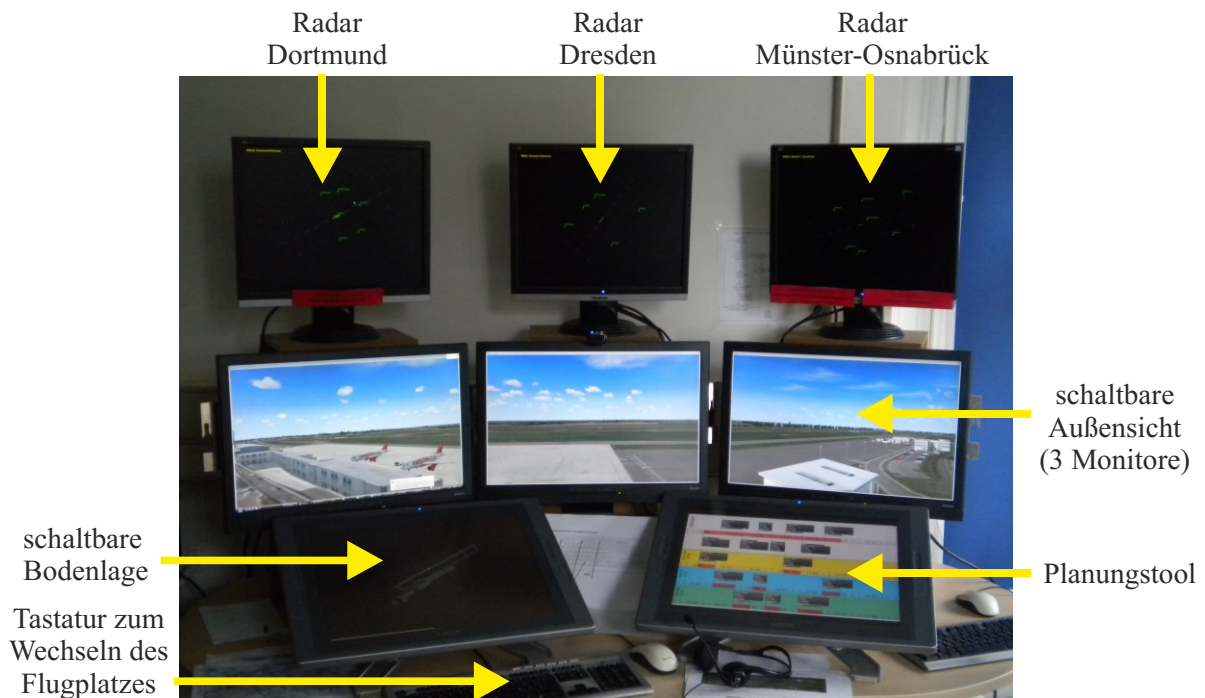


Abbildung 13: Multi Remote Tower-Arbeitsplatz

1.4.3 Mögliche Einschränkungen bei der Nutzung der Echtzeitsimulation

Es wurde in der Einleitung bereits erwähnt, dass der Unfall das Ergebnis einer Verkettung von einzelnen Ursachenfaktoren ist. Die Fähigkeit zur Herstellung derartiger Verkettungen durch die Echtzeitsimulation ist somit entscheidend für den Nutzen als Mittel der Risikoanalyse. Es ist jedoch zu vermuten, dass das Risiko mit Hilfe von Echtzeitsimulationen aufgrund der geringen Stichprobenzahl und der Künstlichkeit alleine nicht bestimmt werden kann (Brooker, 2010) (Shorrock S. T., 2001). Diese Erkenntnis wird durch die beiden Beispiele der simulationsgestützten Risikoanalyse in den Abschnitten 1.4.2.3 und 1.4.2.4 verdeutlicht. Eine treffende Beschreibung der Problematik für die Verwendung von Simulatoren für die Analyse von menschlichen Fehlern lieferte Swain bereits 1990:

„Less-than-adequate calibration of simulator data: because the simulator is not the real world, the problem remains how raw data from training simulators can be modified to reflect real-world performance“ (Swain A. D., 1990, S. 308).

Als ein dazu passendes Beispiel kann das künstliche Verhalten von Menschen in Echtzeitsimulationen genannt werden, das ein gegenwärtig immer noch ungelöstes Problem darstellt. Die Ursache kann z.B. das de facto risikofreie Bedienen des Simulators sein, bei der

der Vorteil des Simulators somit auch gleichzeitig einen Nachteil darstellt. Ebenso können bislang unbekannte Simulations-bedingte Artefakte der künstlichen Arbeitsumgebung Einfluss auf das menschliche Verhalten haben. Die folgenden Abschnitte heben zwei wesentliche Probleme mit einer vertieften Diskussion über die Gründe hervor.

1.4.3.1 Die Künstlichkeit der Arbeitsumgebung

Die für die Risikoanalyse notwendige Realitätstreue der Arbeitsumgebung, die zur Abbildung der Unfallentstehung gefordert wäre, wird einleitend am Beispiel der Analyse von Unfallberichten der Jahre 1979 bis 2009 der Kategorie *Loss Of Control* in Abbildung 14 illustriert (Belcastro & Foster, 2010). Das Ziel der Analyse war die Identifikation von Risikoschwerpunkten in den Ursachen dieser Unfallkategorie. Das Auftreten von

- Systemfehlern alleine oder in Kombination mit Luftfahrzeugstörungen oder
- Schnee/Eis, die zu einer Beeinträchtigung des Luftfahrzeugs führen,

stellen demnach den Risikoschwerpunkt dieser Unfallklasse dar. Das Entdecken von kausalen Mustern gehört nach wie vor zu einer methodisch herausfordernden Disziplin. Derartige Studien verdeutlichen dabei die Komplexität der Unfallentstehung als intensives Zusammenspiel kleinster Elemente, die (wenn bekannt) allenfalls als Faktor auf einer generischen Ebene abgebildet werden können. Die Probleme der Überführung einer derartigen Komplexität in eine simulierte Arbeitsumgebung sind anhand des Induktionsproblems erklären.

Dem Aufbau einer Arbeitsumgebung, die derartige Bedingungen des Unfallherganges unterstützt, stehen zwei grundsätzlich konzeptionelle Schwachstellen der Verwendung der Echtzeitsimulation entgegen:

- *Verallgemeinerbarkeit des Unfalls*: Bei der Rekonstruktion einer Unfallentstehung z.B. im Rahmen einer Unfalluntersuchung ist es eine wiederkehrende Erkenntnis, dass zum Unfallzeitpunkt eine einzigartige Kombination von Ursachenfaktoren vorlag. Der Unfallbeitrag lässt sich aufgrund dessen nur sehr schwer anhand von Faktoren wie z.B. Niederschlagswerten, Höhe der Wolkendecke, Sichtweite und Windstärke und -richtung usw. formalisieren und beschränkt verallgemeinern. Eine Echtzeitsimulation kann eine derartige Verallgemeinerbarkeit (externe Validität) einer Unfallentstehung aufgrund der künstlichen Arbeitsumgebung nicht leisten.

„Die externe Validität sinkt mit wachsender Unnatürlichkeit der Untersuchungsumgebungen...“ (Bortz & Döring, 2006, S. 53)

Es kann aus diesem Zitat geschlossen werden, dass es neben der externen verschiedene Typen der Validität gibt. Die soeben aufgeführten Umweltfaktoren z.B. betreffen vornehmlich die sogenannte *Physikalische Validität* (Mullen, Charlton, Devlin, & Bedard, 2011), welche ein Maß für die Übereinstimmung des Simulatorverhaltens mit dem realen Vorbild darstellt. Die Validität ist insbesondere für die Verallgemeinerung von absoluten numerischen Werten begrenzt, wie z.B. von Wahrscheinlichkeitswerten. Dieser *Absoluten Validität* steht die Gültigkeit von Ergebnissen gegenüber, die den Trend von Effekten beschreiben und sie in einen Vergleich zu anderen Versuchskonfigurationen setzen. Dies wird als *Relative Validität* bezeichnet.

- *Modellunsicherheit*: Die begrenzte Abbildung der Realität ist ein inhärentes Merkmal eines Modells. Dieses systematische Defizit ist die Ursache für die zuvor erläuterte beschränkte Verallgemeinerbarkeit und ist durch sogenannte Modellunsicherheiten begründet. Dies lässt sich mit Hilfe der Mengenlehre erklären. Die Abbildung 15, links, zeigt passend dazu eine Karte, bei der die schrittweise Konkretisierung der Realität zur „relevanten Realität“, zum Modell und vom Modell zum Parameterset dargestellt wird (Pelz & Hedrich, 2015). Die nur teilweise Berücksichtigung der relevanten Realität aufgrund von Simplifikationen und Annahmen durch das Modell führt zu Unsicherheiten in der Vorhersagegenauigkeit des Realverhaltens. Die Unsicherheit kann somit vereinfacht vorgestellt als Lücke zwischen Realität und Modell verstanden werden.

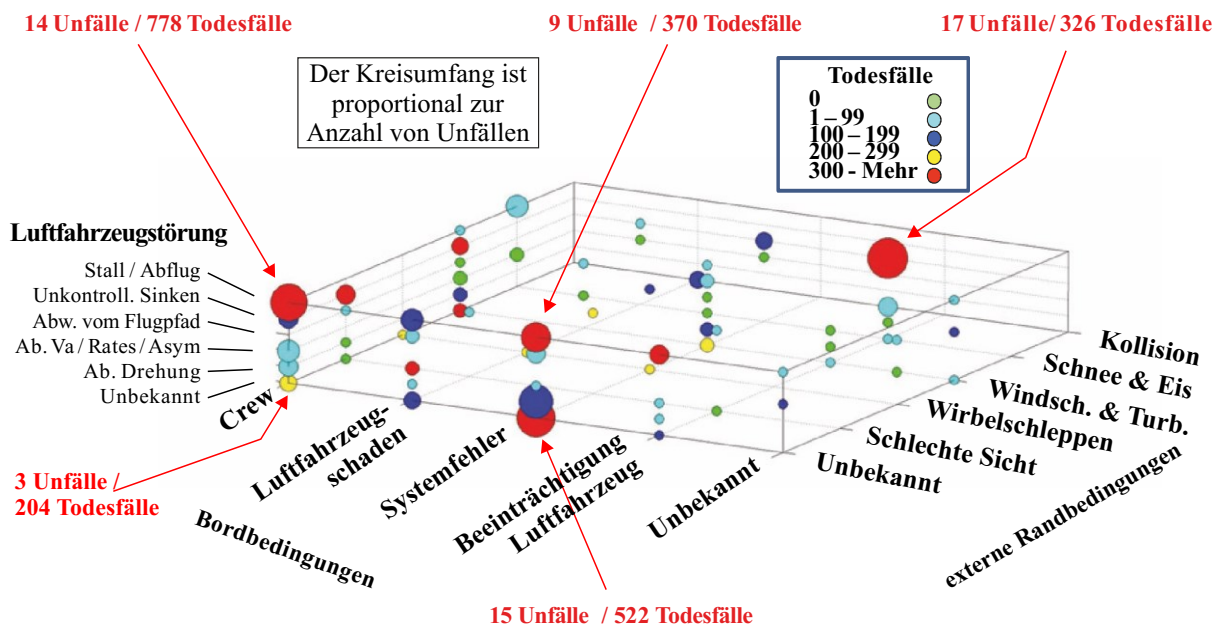


Abbildung 14: Identifikation von Schnittmengen in den Unfallursachen (Belcastro & Foster, 2010)

Der Abgleich zwischen der relevanten und abgedeckten Menge erfolgt i.d.R. anhand von unterstellten Fällen, die vorab bekannt sind. Die A-priori-Eingrenzung einer sicherheitsrelevanten Menge an Faktoren ist nach gegenwärtigem Stand des Wissens allerdings nicht möglich, da nicht alle Gefahren und die dazugehörigen situativen Faktoren im Vorhinein zuverlässig und flächendeckend bestimmt werden können. Dieses Problem wurde bereits einleitend im ersten Abschnitt des Kapitels 1 vorgestellt und als zentrales Problem gegenwärtiger Risikoanalysen herausgestellt. Dieses Problem illustriert die Abbildung 15, rechts, für das allgemeine „Air Navigation System“, welches ein soziotechnisches System ist. Die Abbildung verdeutlicht anhand der Teilmengen die lückenhafte Identifikation von unvorstellbaren Gefahren mit Hilfe der gegenwärtigen Verfahren der Gefahrenidentifikation. Die endlose Suche nach allen Gefahren und den dazugehörigen Faktoren kann als *Analyseparalyse* bezeichnet werden (Brooker, 2010), deren Ergebnis niemals Vollständigkeit erlangen wird. Die Folge dieser Unsicherheit ist eine unbekannte Vollständigkeit des ermittelten Risikos gegenüber den möglichen Wegen der Unfallentstehung.

Für die kontinuierliche Erweiterung und Verbesserung der Modelle und der Simulatoren ist es notwendig, den Einfluss der Künstlichkeit der Arbeitsumgebung weiter zu verringern. Kann bei der Versuchsperson eine natürliche Verhaltensweise in akzeptablen Grenzen ex post festgestellt werden, die durch die Künstlichkeit des Arbeitsumfeldes eine vernachlässigbare Beeinflussung erfuhr, erscheint ein induktiver Schluss möglich. Dies betrifft insbesondere die möglichen Unfallursachen im Menschverhalten.

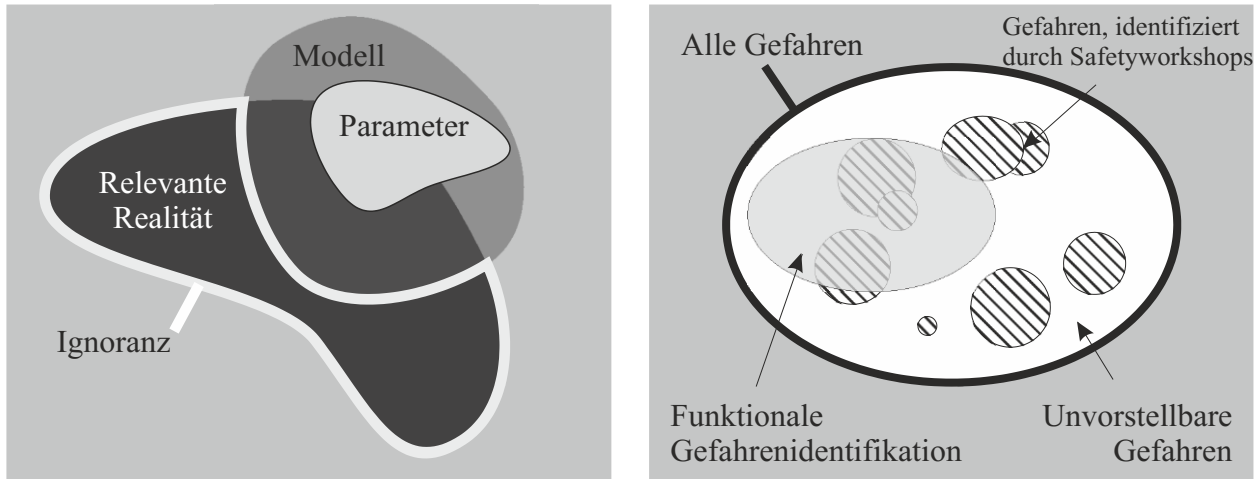


Abbildung 15: Illustration der Modellunsicherheit anhand der Mengenlehre (Pelz & Hedrich, 2015)(links). Gefahren eines Systems (de Jong, 2004) (rechts)

Eine Studie über die Gültigkeit des simulationsgestützten Trainings im Flugbetrieb unter Berücksichtigung der Realitätstreue wurde bislang nicht vorgenommen. Ebenso wurde die unterstellte Physik (z.B. Wetter und Flugverhalten) von CS-FSTD-zertifizierten Flugsimulatoren hinsichtlich ihrer Gültigkeit nie in Gänze untersucht. Beide o.a. Argumente sind Gründe für die beschränkte externe Validität der künstlichen Arbeitsumgebung. Die Möglichkeit eines Induktivschlusses ist aus diesem Grund umstritten.

1.4.3.2 Die Rechtszensur

Die Begrenzung einer verfügbaren Zeit zur Beobachtung eines stochastisch verteilten Zielereignisses kann zu dem Ergebnis führen, dass das Ereignis innerhalb dieser Zeit nicht zu beobachten war. Dieses Problem ist bekannt als statistische Rechtszensur, bei der von der Zeit bis zum Eintritt des Zielereignisses lediglich bekannt sein kann, dass sie größer als die verfügbare Zeit ist (Nelson, 2004). Dies ist insbesondere von Relevanz, wenn sehr kleine Wahrscheinlichkeiten mit geringer Eintrittsfrequenz vorliegen und der Eintritt des Zielereignisses innerhalb der verfügbaren Simulationszeit sehr unwahrscheinlich ist. Bei Simulationsstudien in der Luftfahrt ist es aus Gründen der Mittelverfügbarkeit, der praktischen Organisation und der Projektlaufzeit von Sicherheitsstudien üblich, einige Hundert Stunden zu simulieren (Stroeve, Blom, & Bakker, 2013).

Der erwartete Zeitraum, in dem die Reaktion erfolgt bzw. das Zielereignis eintritt, kann durch die Vorgabe eines stimulierenden Ereignisses eingegrenzt werden. Die Anwendung des prospektiven Ansatzes unterliegt aufgrund der gezielten Vorgabe eines solchen Ereignisses somit nicht diesem Effekt.

Im Gegensatz dazu ist der retrospektive Ansatz für den Fall problematisch, wenn die Wahrscheinlichkeiten sehr viel kleiner sind als die zur Verfügung stehende Simulationszeit. Eine Risikoanalyse würde bei Unfallwahrscheinlichkeiten, die z.B. kleiner als $1e-6$ pro Flugstunde sind, bei einigen Hundert simulierten Stunden wahrscheinlich kein Zielereignis zur Analyse erfassen. Dabei bewirkt der entstehende Effekt rechtszensierter Ereignisse einen logischen Fehlschluss nach Beendigung der Simulationsstudie: die fehlerhaften Schlussfolgerungen über die erfolgreiche Erfüllung von Zielsicherheiten, wenn kein Unfall- oder Vorfalleignis beobachtet werden konnte. Dieser Fehler ist insbesondere der Untersuchungshypothese geschuldet, die bei einem negativen Befund der Studie die Erfüllung der Zielsicherheiten bestätigen soll. In der Praxis ist eine derartige Auslegung der Untersuchungshypothese eine aus der Not geborene „Alternativ“-Lösung, um das Problem der Rechtszensur zu vermeiden. Als aktuelles Beispiel kann hier eine Untersuchung zur Machbarkeit und Validierung von Multi Remote Tower im Rahmen des SESAR WP6.9.3, Versuch VP-061 aufgeführt werden (SESAR, 2015). Da nach 56 Stunden Simulationszeit kein sicherheitsrelevantes Ereignis zu beobachten war, kamen die Untersuchenden zu dem Schluss, dass die Sicherheit durch das neue Betriebskonzept nicht beeinträchtigt sei. Das Problem der Rechtszensur ist somit ein charakteristisches Problem der retrospektiven Risikoanalyse.

1.4.4 Anwendungsszenarien der simulationsgestützten Risikoanalyse

Die Anwendung der Echtzeitsimulation unterstützt nach Betrachtung der Vor- und Nachteile grundsätzliche die Phasen des Safety Case-Projekt-Lebenszyklus, die in Abschnitt 1.3.7 vorgestellt wurden. Wesentliches Kriterium für eine Unterstützung ist hierbei die Gültigkeit der Echtzeitsimulation bezüglich der durch die Simulation generierten und beobachtbaren Ereignisse. Abbildung 16 bietet hierzu eine Übersicht über die möglichen Untersuchungsansätze mit einer Illustration des Gültigkeitsbereichs der simulierten Arbeitsumgebung. Die Untersuchungsrichtung ist im jeweiligen Feld mit einem Pfeil gekennzeichnet. Mögliche Anwendungsszenarien der simulationsgestützten Risikoanalyse können z.B. die Folgenden umfassen:

1. *Verifikation der Gefahrenfolgen*: Insbesondere wenn ein möglicher Ursachenfaktor oder eine spezifische Gefahrensituation als Kombination von möglichen Ursachen hinsichtlich der Folgen geprüft oder die Risikosensitivität bestätigt werden soll, ist der deduktiv-prospektive Ansatz eine geeignete Form der Untersuchung. Untersuchungsziel ist eine Kausalhypothese, die anhand des Eintritts von vordefinierten sicherheitsrelevanten Ereignissen die Gefahr als risiko-induzierend verifiziert. Ein typischer Anwendungsfall ist der „Relative Safety Case“ gemäß (Eurocontrol, 2006a), der anhand sogenannter „Relative Safety Statements“ die Sicherheitsrelevanz von möglichen Gefahren gegenüber einer Baseline falsifizieren oder verifizieren kann. Die Abbildung 16 zeigt hierzu oben rechts einen Ereignisbaum, dessen zu messende Zieleffekte des initiierten Ereignisses innerhalb eines gültigen Bereiches des simulierten Arbeitsumfeldes liegen müssen.
2. *Bestimmung der Gefahrenfolgen*: Der Schweregrad und die damit verbundene Wahrscheinlichkeit sind übliche Parameter für die Ereignisbaumanalyse. Dies entspricht einem induktiv-prospektiven Ansatz. Aufgrund der Annahmen, unter der die

Risikosensitivität gelten soll, habe viele simulationsgestützten Studien eine sehr situationsspezifisch zugeschnittene Untersuchungshypothese. Als Beispiel kann hier die Studie (Stroeve, Blom, & Bakker, 2013) genannt werden, die eine typische Runway Incursion-Situation standardisiert und einzelne situative Faktoren hinsichtlich ihres Risikos untersucht. Die Abbildung 16 illustriert hierzu oben links einen Ereignisbaum, dessen Gültigkeit der Parameter durch die fehlende Gültigkeit der Simulationsbedingungen verletzt werden kann.

3. *Verifikation von erwarteten Risikoschwerpunkten:* Ist es das Ziel, eine Menge vorab unterstellter Ursachenfaktoren in einer Gegenüberstellung hinsichtlich der Risikosensitivität relativ zu gewichten (Ranking), so ist der deduktiv-retrospektive Ansatz zu wählen (Abbildung 16, unten rechts). Vermutete Risikoschwerpunkte können anhand der Gegenüberstellung der bedingten Wahrscheinlichkeitswerte verifiziert oder falsifiziert werden. Der Ansatz bietet sich z.B. an, wenn Gestaltungsoptionen im Konzept oder Design des Systems hinsichtlich ihrer Nutzbarkeit bewertet werden sollen (usability studies). Ebenso können mögliche Risikoschwerpunkte auf den Betriebsflächen und der Verfahrensplanung Gegenstand der Analyse sein. Als Beispiel kann hier (Meyer L. , Schultz, Schmidt-Roßleben, & Fricke, 2013) genannt werden, in der die fehlerhafte Wahrnehmung visueller Informationselemente des Fluglotsen hinsichtlich ihrer zugehörigen Risikosensitivität geprüft wird. Ein Risikoschwerpunkt konnte hierbei systematisch für den Aufrollbereich der Start- und Landebahnen des Flughafens Frankfurt a. Main identifiziert werden.
4. *Verifikation einer Zielsicherheit:* Für die Verifikation der Erfüllung einer Zielsicherheit z.B. Unfallwahrscheinlichkeit durch die Gesamtheit der bekannten Ursachen kann der deduktiv-retrospektive Ansatz dienen. Dieses Szenario bietet sich für die Erbringung des Sicherheitsnachweises an, wie in der Einleitung in Abschnitt 1.1 angesprochen wurde. Die Bestimmung des Risikos und die dazugehörige Varianz sind für die Verifikation zu berücksichtigen. Die Verifikation ist dann nur für die jeweilig einbezogenen Ursachen gültig, die durch die simulierte Arbeitsumgebung berücksichtigt werden kann.
5. *Identifikation von unbekanntem Risikoschwerpunkten:* Die Änderung von Arbeitsabläufen des Operators kann eine Folge der Weiterentwicklung oder einer Modifikation des Konzepts oder des Designs von Verfahren, Instrumenten oder einer Organisation sein. Eine solche Änderung kann innovationsinduzierte Gefahren hervorrufen, die aus einem bis dahin unbekanntem Zusammenspiel der situativen Faktoren entstehen. Dieser Typ von unvorstellbaren Gefahren wurden in der Abbildung 15, rechts, gegenüber anderen Gefahrentypen abgegrenzt. In diesem Fall wäre ein induktiv-retrospektiver Untersuchungsansatz geeignet, da dieser ein Minimum an notwendigem A-priori-Wissen über Ursachen und deren Risikoschwerpunkte voraussetzt (Abbildung 16, unten links). Das Ergebnis der Schlussfolgerung über mögliche Ursachen und die Risikoschwerpunkte dient zur ergebnisoffenen Ergründung der wichtigsten Gefahren, die es weiter zu analysieren und deren Risiko es zu bestimmen gilt. Insbesondere in der ersten Phase, der Konzeptentwicklung und Implementierung eines Systems, kann hier konkretes und substantielles Wissen über

die Sicherheitsrelevanz von Instrumenten, den Verfahren und dem Training gewonnen werden, die einer unvoreingenommenen Bewertung aller Elemente des soziotechnischen Systems dienen können.

6. *Bestimmung eines unbedingten Risikos*: Die Beobachtung von Unfällen oder anderen sicherheitsrelevanten Ereignissen kann für die Ermittlung eines Gesamtrisikos verwendet werden, welches nicht den Eintritt eines Gefahrenereignisses voraussetzt. Aus diesem Grund wird es als „unbedingt“ bezeichnet. Eine solche Bestimmung ist bei Anwendung des induktiv-retrospektiven Ansatzes möglich. Es ist dabei zu beachten, dass die in Abschnitt 1.4.3.1 vorgestellten Einschränkungen zu einer Ungültigkeit dieses Ergebnisses führen können, da eine absolute Validität durch die Simulationsumgebung prinzipiell nicht geleistet werden kann.

1.5 Die Zielsetzung und Struktur der Arbeit

1.5.1 Die Problemdefinition und der Lösungsansatz

Aufgrund der Notwendigkeit zur Unterstützung und Entwicklung soziotechnischer Systeme in der Luftfahrt hat diese Arbeit die Weiterentwicklung einer Methode der Risikoanalyse zum Gegenstand. Die Methode soll sich auf den retrospektiven Untersuchungsansatz stützen, da dieser jenen Wissensgewinn verspricht, der zu den Phasen des Safety Case-Lebenszyklus Konzept, Implementierung und Migration am notwendigsten ist. Der Vergleich der Anwendungsszenarien des vorhergehenden Abschnitts mit den Problemen der gegenwärtigen Risikoanalyse, vorgestellt in Abschnitt 1.3.6, lässt ebenso hier einen bedeutenden Gewinn für den Nutzen der Echtzeitsimulationstechnik zur Risikoanalyse bei eben diesem Ansatz erwarten. Dieses betrifft bei neuartigen soziotechnischen Systemen insbesondere

- die Verifikation einer Zielsicherheit,
- die Identifikation von unbekanntem Risikoschwerpunkten und
- die Verifikation von erwarteten Risikoschwerpunkten.

Die beiden in Abschnitt 1.4.3 aufgeführten Problemfelder heben die besondere Relevanz der Rechtszensur für den retrospektiven Untersuchungsansatz hervor. Sie wird als zentrales Problem dieser Arbeit erachtet. Bei der Wahl eines geeigneten Lösungsansatzes lohnt sich der Vergleich mit der ähnlichen Problematik zur Schätzung der Lebenszeit technischer Produkte. Ziel dieser Methode ist es, den erwarteten Ausfallzeitpunkt der sogenannten Spätausfallphase (*wear-out failure period*) und seine stochastische Verteilung in Folge von Alterungsprozessen durch z.B. thermischem und mechanischem Stress vorherzusagen. Ebenso sind Prognosen für lange Zeiträume lieferbar, die die erzielte Qualität eines Produktes bewerten können und ermöglichen, diese nach Anforderungen ggf. robuster auszulegen.

In diesem Bereich wurde eine Methode zum beschleunigten Lebensdauertest namens *Accelerated Life Testing* (ALT) entwickelt (Nelson, 2004). Ein Beispiel zur Anwendung dieser Methode wurde bereits unter dem Begriff der beschleunigten Materialermüdungstests vorgestellt. Diese Methode nutzt experimentell gewonnene Daten, die bei Betrieb des Produktes unter modifizierten Realbedingungen gewonnen werden können. Die Modifikation besteht aus der Intensivierung vordefinierter Stressfaktoren, die den Alterungsprozess beschleunigen. Die Kalibrierung der Intensivierung schafft die Voraussetzung, die Größe des wirkenden Stresses aus einem bekannten „Normal“ kontrolliert zu ändern und Stresslasten so

gezielt zu erzeugen. So gelingt es, die statistische Verteilung des Fehlers, der zu dem Ausfall des Testgegenstandes führt, auf der Zeitachse kontrolliert vorzubewegen.

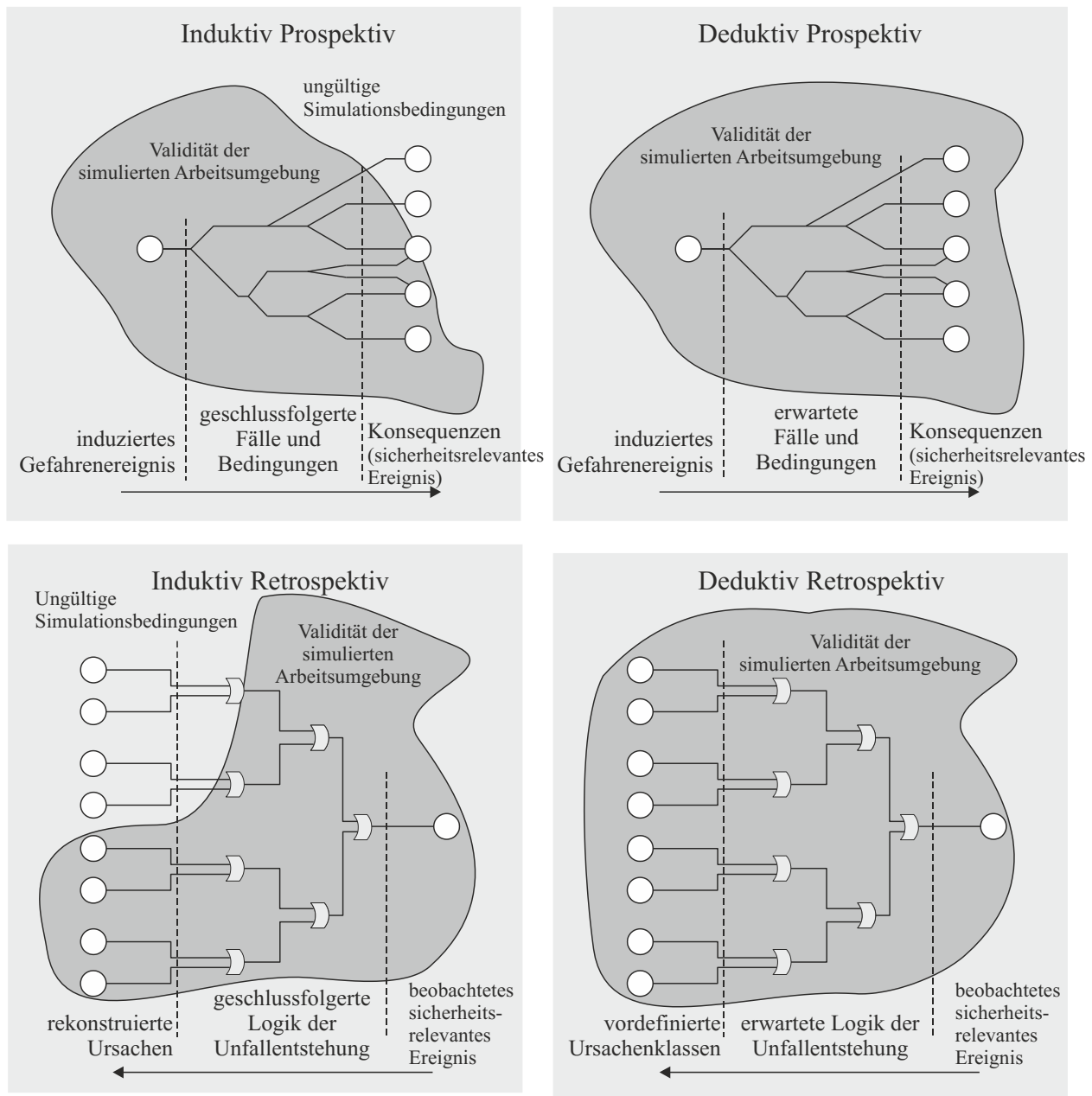


Abbildung 16: Gültigkeit der simulierten Arbeitsumgebung bei Anwendung der Untersuchungsansätze

Der Rückschluss auf die Zielverteilung unter Designstress gelingt schließlich anhand einer Regressionsanalyse, die eine bekannte Stress-Lebenszeit-Relation als Regressionsmodell unterstellt. Die Abbildung 17 liefert hierzu ein Beispiel über 3 Stützstellen der Wahrscheinlichkeitsdichte des Fehlers, die den Rückschluss auf die Zielverteilung erlaubt. Als Beispiel eines bekannten Regressionsmodelles ist die *Arrhenius*-Gleichung zu nennen, die den temperaturbedingten Ausfall eines Geräts in eine exponentielle Abhängigkeit mit der vorherrschenden Temperatur setzt (Condra, 2001). Die Modifikation für eine Stressintensivierung bei technischen Produkten ist häufig durch eine gleichbleibende Systematik des Alterungsprozesses gekennzeichnet. Die Reaktionen der Verteilungen auf die

Stresseinwirkung erscheinen reproduzierbar und die ermittelbare Lebenszeit in die Praxis übertragbar.

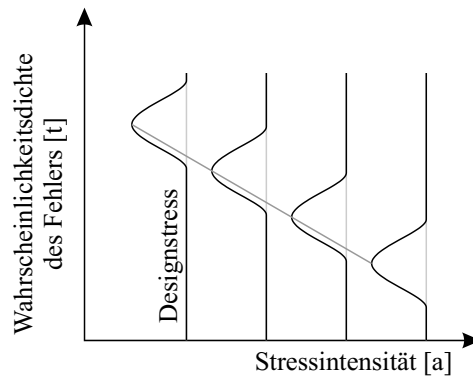


Abbildung 17: Die Verteilung des Ausfallereignisses in Relation zur Stressintensität nach ALT (Nelson, 2004)

1.5.2 Die Untersuchungshypothese

Die Übertragung dieses Konzeptes für die Anwendung in der Risikoanalyse soziotechnischer Systeme ist der untersuchte Lösungsvorschlag dieser Arbeit, der statt des Ausfallereignisses das Unfallereignis als Zielmetrik für die Risikoanalyse setzt. Die zu untersuchende Fragestellung dieser Arbeit leitet sich aus der Vermutung ab, menschliche Fehler unter Einwirkung ähnlicher Stressinduktionsverfahren gezielt stimulieren zu können und somit die Wahrscheinlichkeit seltener sicherheitsrelevanter Ereignisse für die Lösung der Rechtszensur erhöhen zu können. Die Untersuchungshypothese bezieht sich somit auf den Kern der Übertragbarkeit von ALT in den Flugbetrieb: Die Wirkung eines induzierten erhöhten Stress kann die Sicherheitsmarge des Flugbetriebes gezielt und kontrolliert verknappen. Die Sicherheitsmarge kann dabei nach Vorbild Amalbertis (Amalberti, 2001) als ein vorgehaltener Spielraum verstanden werden, der sich an der Differenz zwischen der aktuellen Unfallwahrscheinlichkeit und der Zielsicherheit bemessen lässt.

Die „Sichtbarkeit“ des Unfalls oder eines vordefinierten Vorfallereignisses innerhalb des Experiments soll als messbare Zielgröße die Grundlage zur Einschätzung des Risikos liefern. Die Kenntnis der Stress-Risiko-Relation als Regressionsmodell voraussetzend, wird der Rückschluss von den beobachteten Ereignisdaten auf das Risiko des nicht-modifizierten Normalbetriebs möglich. Ein soziotechnisches System könnte sodann hinsichtlich der Erfüllung einer vorgegebenen Zielsicherheit verifiziert werden.

Der validen Übertragung des Konzeptes ALT für die Risikoanalyse von soziotechnischen Systemen steht das bekanntermaßen vielschichtige und unsichere Menschverhalten gegenüber. Gegenüber technischen Produkten zeichnet es sich vermutlich durch ein begrenzt reproduzierbares Verhalten und eine wenig kontrollierbare Stressreaktion aus. Die Stochastik des menschlichen Verhaltens bildet somit die größte Herausforderung für die Übertragbarkeit. Die Untersuchung, ob eine gezielte und kontrollierbare Intensivierung der Menschfehlerwahrscheinlichkeit gelingen kann, und die Bedingungen, die dazu notwendig sind, sind sodann die ersten Schritte der Übertragung von ALT. Der mögliche Erfolg wird dabei gestützt durch die in Abschnitt 1.2.2 angenommene Korrelation der Unfallwahrscheinlichkeit mit einwirkenden Stressfaktoren. Zur Begrenzung des

Untersuchungsumfangs wird das Anwendungsszenario „Verifikation einer Zielsicherheit“ als exemplarischer Untersuchungsgegenstand gewählt, das anhand der Irrtumswahrscheinlichkeit des Verifikationsergebnisses ein einfaches und gleichzeitig repräsentatives Maß der Rechtszensur bietet.

1.5.3 Die Struktur der Arbeit

Die Gliederung dieser Arbeit orientiert sich an dem Ziel, eine Aussage über die Übertragbarkeit des ALT-Konzeptes auf den Menschen mit Hilfe der entwickelten Methodik *Accelerated Risk Analysis* zu liefern. Die Arbeit gliedert sich zu diesem Zweck in 4 Kapitel.

Die Methodik *Accelerated Risk Analysis* wird in Kapitel 2 vorgestellt. Sie liefert auf Basis des retrospektiven Untersuchungsansatzes eine Lösung für das Problem rechtszensierter Ereignisdaten. Aufgrund der Wahl des Anwendungsszenarios „Verifikation einer Zielsicherheit“ wird die Lösung am Beispiel eines deduktiven Untersuchungsansatzes entwickelt. Die Wahl dieses Ansatzes bietet den Vorteil, Einschränkungen der Validität seitens des Simulators für die in dieser Arbeit vorgenommenen Untersuchungen ausschließen zu können. Die Probleme des Induktivschlusses werden aus diesem Grund als unabhängig zu untersuchendes Problem erachtet.

Die Methode umfasst ein Risikomodell, welches Stressinduktion, Stressreaktion und Unfallwahrscheinlichkeit kausal verknüpft und den zu untersuchenden Zusammenhang für den beschleunigten Stresstest liefert. Ebenfalls wird in diesem Kapitel das Stressinduktionsverfahren erläutert, welches den Menschen (Probanden) in eine Wettbewerbssituation versetzt und somit zeitdruckintensivierte Arbeitsbedingungen herstellt. Begleitend werden die Zielmetriken und Verteilungen definiert, die für das Konzept begründet unterstellt werden.

Das darauf folgende Kapitel „Experimentelle Studie“ soll das Prinzip testen, mit Hilfe des Zeitdrucks die Unfallwahrscheinlichkeit zu intensivieren und die Sicherheitsmargen gezielt zu senken. Dies stellt eine Teilvalidierung des Konzeptes *Accelerated Risk Analysis* dar, dessen Ergebnisse Aussagen über die Effektivität des Zeitdruckinduktionsverfahrens zulassen. Das Kapitel umfasst sowohl das Experimentaldesign mit der Entwicklung einer geeigneten Untersuchungsumgebung am Beispiel der Flugplatzkontrolle des Flughafens Frankfurt a. Main, die Durchführung und die Auswertung hinsichtlich der Stressreaktionen und die gemessene Anzahl an sicherheitsrelevanten Ereignissen. Die Implementierung des Stressinduktionsverfahrens beruht auf einem softwarebasierten Fluglotsenagenten, dessen Arbeitsgeschwindigkeit den Wettbewerbsdruck mit einem parallel hierzu agierenden Probanden für diesen im Betrieb aufbaut. Anhand von zahlreichen Aufzeichnungsdaten können die Stressreaktionen der Probanden nachvollzogen werden.

Das abschließende Kapitel 4 beinhaltet ein Urteil über die Wirksamkeit und die Praktikabilität des Konzeptes für die Risikoanalyse. Anhand von Erklärungsansätzen dritter werden die Stressreaktionen diskutiert, um das Verhalten der Probanden zu verstehen und um weitere Entwicklungsmöglichkeiten des Konzeptes aufzuzeigen.

2 Konzeptentwicklung „Accelerated Risk Analysis“

Wie in der Einführung, Abschnitt 1.3, aufgeführt, ist die Qualität einer simulationsgestützten Risikoanalyse aufgrund der Komplexität der Unfallentstehung als auch des Effektes der Rechtszensur wesentlich beeinträchtigt. Das in diesem Kapitel vorgestellte Konzept *Accelerated Risk Analysis* (AccSis) hat vor diesem Hintergrund den Zweck, von den zwei aufgeführten Problemfeldern einen Beitrag zur Lösung des Problems rechtzensurierter Daten zu leisten. Das Konzept AccSis beschreibt dazu eine Methode, die das Beobachtbarkeitsproblem von Unfallereignissen mit kleinen Wahrscheinlichkeiten bei begrenzt zur Verfügung stehender Simulationszeit explizit adressiert. Sie soll es gestatten, die Erfüllung von Zielsicherheiten (*Target Level of Safety*, TLS) neuartiger soziotechnischer Systeme in der Luftfahrt mit einer akzeptablen Irrtumswahrscheinlichkeit experimentell verifizieren zu können.

Der retrospektive Untersuchungsansatz ist aufgrund der Unabhängigkeit von einem initialen Gefahrenereignis die gewählte Lösung, um das Risiko neuartiger soziotechnischer Systeme analysieren zu können. Das dahinter stehende Prinzip dieses Ansatzes wurde bereits im ersten Abschnitt von Kapitel 1 argumentiert und in Abschnitt 1.4.4 anhand möglicher retrospektiver Anwendungsfälle beschrieben.

Die Irrtumswahrscheinlichkeit, den falschen Schluss über das Erfüllen einer Zielsicherheit zu ziehen, wird vor diesem Hintergrund als ein Maß zur Indikation des Effektes der Rechtszensur erachtet. Die vorgeschlagene Lösung zur Minderung dieses Effektes besteht aus der Herabsetzung der vorgehaltenen Sicherheitsmargen des simulierten Flugbetriebs durch eine gezielte Intensivierung der Unfallwahrscheinlichkeit. Der stochastische Effekt, der durch die Intensivierung zum Tragen kommt, ist eine schnellere Annäherung der beobachteten Unfallhäufigkeit an die erwartete. Dies ist durch die *stochastische Konvergenz* erklärbar, die die Konvergenz des Mittels der Differenz zwischen einer Folge von Zufallszahlen und der Wahrscheinlichkeit der dazugehörigen Grundgesamtheit bei zunehmender Stichprobenzahl beschreibt (Sachs, 1999). Wird der Unfall als stochastisches Zufallsereignis erachtet, beschreibt *das schwache Gesetz der großen Zahlen* eine eben solche beschleunigte Konvergenz bei Erhöhung der Unfallwahrscheinlichkeit. Eine kleinere Differenz bei gleicher Simulationszeit soll schließlich zu einer kleineren Irrtumswahrscheinlichkeit führen, wenn aus der beobachteten Unfallhäufigkeit auf die Unfallwahrscheinlichkeit geschlossen werden soll.

Die vorgestellte Problemanalyse und Lösungsentwicklung dieses Kapitels nimmt häufig Bezug auf die Begriffe *Wahrscheinlichkeit* und *Häufigkeit*. Die Verbindung zwischen den beiden Begriffen ist für diese Arbeit wie folgt gewählt: Die Wahrscheinlichkeit ist als Erwartungswert eine relative Häufigkeit¹², um erwartete Häufigkeiten vorherzusagen (Sachs, 1999). Der Begriff der *Rate* ist ähnlich, die sich allerdings auf eine beobachtete relative Häufigkeit bezieht.

Für die Entwicklung der Methode, die der anschließenden Verifikation einer Zielsicherheit dienen soll, steht, wie in Kapitel 1 dargelegt, die Methode *Accelerated Life*

¹² Die relative Häufigkeit wird hier allgemein als das Verhältnis zwischen der Anzahl eines Ereignisses und der Anzahl der Gelegenheiten für den Eintritt des Ereignisses erachtet. Demgegenüber steht die absolute Häufigkeit, die Ausdruck für die Anzahl eines Ereignisses ist.

Testing (Nelson, 2004) Vorbild. Der Aufbau von AccSis orientiert sich aus diesem Grund an den Verfahrensschritten dieser Methode:

1. Fehlerstimulation,
2. Regression der Lebenszeitverteilung der jeweiligen Stresslasten und
3. Regression des Lebenszeit-Stress-Zusammenhangs.

Hierauf aufbauend werden die folgenden äquivalenten Verfahrensschritte für AccSis vorgestellt:

1. *Stimulation menschlicher Fehler*: Für die Steuerung der Intensivierung wurde einwirkender Stress ausgewählt, der menschliche Fehler gezielt herbeiführen soll. Der Abschnitt 1.2.1 hat den menschlichen Fehler als eine bekannte und gut erforschte Unfallursache bereits herausgestellt. Ebenso wurde in diesem Abschnitt der wesentliche Beitrag durch Stress hervorgehoben. Die Induktion von Stress kann demnach für die Intensivierung und somit für die gezielte Herabsetzung der Sicherheitsmargen im Flugbetrieb benutzt werden.

Als vielversprechendster Stressor fällt die Wahl nach Recherche auf den Zeitdruck, der beim Menschen emotionale Stresszustände hervorzurufen im Stande ist. Die Diskussion zur Auswahl des Zeitdrucks erfolgt im später folgenden Abschnitt 2.5.2. Ein technisches Verfahren zur Induktion des Stresses wird vorgestellt, das Zeitdruck in dosiertem Maße während der Echtzeitsimulation aufbauen soll. Es soll dem Operateur das Gefühl vermitteln, die Zeit für die Entscheidungen sei in einem vorgegebenen Maß begrenzt. Das dazu entwickelte Verfahren namens *Konkurrenzdruck* soll die Qualität menschlicher Arbeit mittels Zeitdruckinduktion mindern und als Steuerung für die Intensivierung der Unfallwahrscheinlichkeit dienen. Die umfassende Verknüpfung der Stresseinwirkung mit der daraus entstehenden intensivierten Unfallwahrscheinlichkeit soll schließlich mit Hilfe eines kausalen Risikomodells beschrieben werden, das die einzelnen Prozessschritte und die dazugehörigen Effekte miteinander verbindet.

2. *Bestimmung der Parameter der Unfallverteilung*: Fortführend kommt eine Regressionsanalyse zur Anwendung, deren Regressionsmodell die Verteilung des Unfallereignisses als Funktion einer induzierten Stresslast beschreibt. Zweck dieses Modells ist die Beschreibung der Intensivierung gegenüber dem „Normal“.
3. *Verifikation der Zielsicherheit*: Der Rückschluss von den intensivierten Unfallhäufigkeiten zu nicht-modifizierten „normalen“ Stresslast-Bedingungen (*Design Stress*) soll schließlich unter Zuhilfenahme eines Regressionsmodells, der sogenannten Stress-Risiko-Relation, gelingen. Die Regressionsanalyse setzte dabei voraus, über hinreichend Stützstellen zu verfügen.

2.1 Die grundlegenden Begriffe

2.1.1 Die Wahl der Risikometrik

Im prospektiven Sinne ist das Risiko der Ausdruck für die Wahrscheinlichkeit, mit der ein Schaden aus einer Gefahrensituation entstehen kann. Wird der Eintritt eines Initialereignisses, der Gefahr, nicht unterstellt (retrospektiv), so stellt der Begriff Risiko einen Zusammenhang zwischen dem Schweregrad und seiner Häufigkeit pro Zeiteinheit her. Diese Definition stellt sicher, dass häufige kleine Schäden und seltene große Schäden dem gleichen Risiko

zugeordnet werden können. Der Freiheitsgrad, den diese Definition zulässt, wird z.B. für die Definition der Risikomatrix mit zugehörigen Zielsicherheiten genutzt.

Die Organisation ICAO erachtet das Unfallereignis und seine Vorfalleereignisse als oberste Schweregrade einer Konsequenz und setzt sie somit zur absoluten Maßgröße des Risikos der operationellen Sicherheit:

„...the elimination of aircraft accidents and/or serious incidents remains the ultimate goal...“ (ICAO, 2013c).

Die Relation zwischen dem Unfallereignis und seiner Wahrscheinlichkeit ist nach Vorbild (Blom & Bakker, 2003) eine übliche Risikometrik, die wie folgt definiert sein kann: Das Risiko ist die Wahrscheinlichkeit für den Eintritt eines Unfallereignisses innerhalb eines gegebenen Zeitintervalls.

Für die Auffassung der Begriffs „Unfall“ wird für eine genau Definition auf die Verordnung der *Europäischen Union 966/2010* (The European Commission, 2010) verwiesen, die für die Untersuchungen von Flugunfällen in allen Mitgliedsstaaten der EASA gilt. Im Bereich *Air Traffic Management* (ATM) kann das Unfallereignis gemäß *Integrated Risk Picture* (Eurocontrol, 2006b) beispielhaft nach fünf Typen unterschieden werden:

- Zusammenstoß in der Luft (*Mid-air Collision*),
- Kollision auf der Start- und Landebahn (*Runway Collision*),
- Rollwegkollision (*Taxiway Collision*),
- ungewollte Bodenberührung (*Controlled Flight into Terrain*) und
- Wirbelschleppen-bedingter Unfall (*Wake Turbulence Accident*).

Gemeinsame Schnittmenge aller Typen ist ein Unfall mit der Bezeichnung *Fatal Accident*, der den Tod wenigstens eines Besatzungsmitgliedes oder eines Passagiers innerhalb von 30 Tagen zur Folge hat (ICAO, 2010). Ereignisse, die diesen Klassen zugeordnet werden können, gelten innerhalb dieser Arbeit als primäre Risikometrik zur Bestimmung des Risikos.

Stehen Ereignisse in einer ursächlichen Wirkungsweise zu dem Unfallereignis, so spricht man von einer Vorfall (*Incident*). Dieser Ereignistyp kann für eine Risikobestimmung ebenso angewandt werden (The European Commission, 2010). Die mit diesem Vorfalleereignis verbundene Unfallwahrscheinlichkeit kann dann im prospektiven Sinne anhand einer bedingten Wahrscheinlichkeit beschrieben werden. Der indirekte Schluss von einer Häufigkeit von Vorfalleereignissen pro Zeit auf das Risiko gelingt dann mit Hilfe dieser sekundären Risikometrik.

Einen derartigen Zusammenhang unterstellt z.B. die Unfallpyramide nach *Heinrichs Gesetz* (Heinrich, Peterson, & Roos, 1931), die zwischen Vorfalleereignissen nach Schadensausmaß abgestuft und dem Unfallereignis eine Korrelation mit konstanten Verhältnissen feststellen konnte, wie z.B. 29 leichte Unfälle auf einen schweren Unfall. Die Proportionen dieser Zusammenhänge können allerdings je nach Definition und Auslegung der Schweregrade stark variieren. Diesem Ansatz folgt ebenso das *Accident-Incident-Model* (Eurocontrol, 2006b), das bereits in Abschnitt 1.3.5 vorgestellt wurde.

Ein derartiger Zusammenhang trifft z.B. auf das Ereignis *Runway Incursion*¹³ zu, das anhand von Vorfalldaten auf ein Verhältnis von 1:100 zum Unfalleintritt bestimmt werden konnte (Eurocontrol, 2010b). Dies bedeutet ein Unfall in 3,7 Jahren, der im direkten Zusammenhang mit einer Runway Incursion steht. Andere Vorfalleignisse wie z.B. der Staffelungsverlust (*Loss of separation*) konnten ebenso nach dem soeben genannten Beispiel als unfallursächlich eingestuft werden (Eurocontrol, 2006b).

2.1.2 Die Wahl der Zielsicherheit

Der Begriff der Zielsicherheit (eng. *Target Level of Safety*) drückt die Anforderung an ein System aus, ein vordefiniertes Maß an Schaden pro Zeit, Gelegenheit oder Ereignis nicht zu übersteigen. Sie ist gemäß Eurocontrol's *ESARR 4* eine maximale tolerierbare Wahrscheinlichkeit¹⁴ und der Ausdruck eines akzeptablen Risikos für ATM (Eurocontrol, 2001). Sie ist schnell zu verwechseln mit dem sogenannten *Safety Objective*, das sich gemäß *ESARR 4* explizit auf die maximale Eintrittswahrscheinlichkeit von Gefahrenereignisse selbst bezieht.

Die Zielsicherheit ist erfüllt, wenn die erwartete relative Häufigkeit für den Eintritt eines Unfallereignisses p kleiner als oder gleich einer gegebenen Zielsicherheit p_0 gemäß

$$p \leq p_0 \quad (2.1)$$

ist.

2.1.3 Die Wahl der Grundgesamtheit

Die Menge der Stichprobe stellt einen Auszug der Grundgesamtheit dar, bei der die jeweilige Stichprobe eine Gelegenheit für den Eintritt des Unfallereignisses darstellt. Innerhalb der Menge der Stichproben führt eine bestimmte Anzahl zu dem tatsächlichen Eintritt dieses Unfallereignisses und begründet anhand der Relation zur Menge der Stichprobe die *Wahrscheinlichkeit für den Eintritt des Unfallereignisses* wie folgt

$$p = \frac{\text{Erwartete Häufigkeit einzutretender Unfallereignisse}}{\text{Häufigkeit der Gelegenheit für den Eintritt eines Unfallereignisses}} \quad (2.2)$$

Durch die Wechselwirkung zwischen Entscheidungssituationen und dem resultierenden Betriebsablauf im operativen Umfeld ist die Definition einer statistischen Einheit zum Zwecke der Abgrenzung von unterschiedlichen Gelegenheiten in der Praxis schwierig. Dies kann mit der fehlenden Unabhängigkeit einer Einzelhandlung der Operateure im Flugbetrieb begründet werden. Die Definition einer unabhängigen Gelegenheit ist aufgrund der Vielfältigkeit und der Verfolgung übergeordneter Ziele, mit der die vorliegende Betriebssituation aus Sicht der Flugsicherung als auch der Luftfahrzeugführer geführt wird, impraktikabel und vermutlich wenig zutreffend für die Beschreibung der Realität.

Der Flugbetrieb folgt in der Praxis einem kontinuierlichen Vorgang von Entscheidungsfindungen. Der Begriff „Entscheidungskontinuum“ trifft dieses Merkmal recht

¹³ „Any occurrence at an aerodrome involving the incorrect presence of an aircraft vehicle or person on the protected area of a surface designated for the landing and take-off of aircraft“ (ICAO, 2007b).

¹⁴ „Maximum tolerable probability“ (Eurocontrol, 2001, S. 16).

gut und steht der Definition einer statistischen Einheit konzeptionell entgegen. Er beschreibt, dass eine einzelne Entscheidung oder Handlung ein Teil eines fortdauernden Gesamtführungsprozesses in der Zusammenwirkung verschiedener Operateure ist.

Eine alternative Lösung bietet die Definition einer statistischen Einheit mit einer nicht näher zu bestimmenden Anzahl an Gelegenheiten für einen Unfall, wie z.B. pro Flugstunde. Die Anzahl an Gelegenheiten pro Flugstunde weist dabei eine Varianz auf, die als annähernd konstant erachtet wird. Der Nachteil ist die Vernachlässigung der Varianz der Unfallwahrscheinlichkeit innerhalb der Zeiteinheit. Dieses konzeptionelle Bindeglied stellt somit die Verbindung zwischen der Gelegenheits-basierten Unfallwahrscheinlichkeit und dem Risiko her. Diese Arbeit unterstellt an dieser Stelle die Flugstunde als statistische Einheit der Grundgesamtheit, die zu einer äquivalenten Nutzung der Begriffe Unfallwahrscheinlichkeit und Risiko führt.

2.1.4 Die Wahl der Unfallereignisverteilung

Bei Verwendung der Flugstunde als Grundgesamtheit kann unterstellt werden, dass die maximale Anzahl an aufzutretenden Unfällen innerhalb einer Stichprobe null oder eins sein kann. Dies unterstellt, dass pro Flugstunden nicht zwei oder mehr Unfälle auftreten können, was in Europa eine gültige Annahme für die Risikoanalyse ist. Dies entspricht der Verteilungsbeschreibung des *Bernoulli-Experiments*, in der die Flugstunde die zwei Zustände „Erfolg“ und „Nicht-Erfolg“ hat. Bei der Zusammenfassung mehrerer Bernoulli-Experimente mit der Anzahl n mit der Erfolgsrate k entsteht die Binomial-Verteilung n über k , deren Gültigkeit für die Verteilung des Unfallereignisses über mehrere Flugstunden oder Bewegungen unterstellt wird.

2.2 Das Sicherheitsargument

Die Schätzung einer Unfallwahrscheinlichkeit p auf Grundlage einer gemessenen *Häufigkeit des Unfalls* X_n (bzw. bei Berücksichtigung des Bezugszeitraums dessen *relative Häufigkeit* \bar{X}_n) weist mit Hilfe des Schätzers $\bar{X}_n \approx p$ und bei begrenzter Anzahl an *Stichproben* n grundsätzlich einen Fehler auf. Dieser ist definiert als die Differenz zwischen beobachteter relativer Häufigkeit und der Wahrscheinlichkeit $|\bar{X}_n - p| \leq \epsilon$ mit ϵ als ein gegebener *Toleranzwert*. Nimmt man an, es gäbe für die Beweisführung über die Erfüllung einer Zielsicherheit p_0 eine Versuchsanordnung mit einer begrenzten Anzahl von n Stichproben (Zeiteinheiten oder Handlungssituationen), so sind zwei Fälle und deren Irrtumswahrscheinlichkeiten entscheidend für die Sicherheit der Aussage:

1. Die Irrtumswahrscheinlichkeit, ein unsicheres System fälschlicherweise als sicher zu deklarieren, wenn die beobachtete relative Häufigkeit die Zielsicherheit erfüllt:

$\bar{X}_n \leq p_0$. Diese Irrtumswahrscheinlichkeit ist dann definiert als

$$P(p > p_0 \wedge \bar{X}_n \leq p_0).$$

2. Die Irrtumswahrscheinlichkeit, ein sicheres System fälschlicherweise als unsicher zu deklarieren, wenn die beobachtete relative Häufigkeit die Zielsicherheit nicht erfüllt: $\bar{X}_n > p_0$. Diese Irrtumswahrscheinlichkeit ist dann definiert als

$$P(p \leq p_0 \wedge \bar{X}_n > p_0).$$

Die Theorie der Hypothesenprüfung bezeichnet die zugehörigen statistischen Fehlerfälle als 1. Art und 2. Art. Bei der Zuordnung zu den o.g. Irrtumswahrscheinlichkeiten ist es entscheidend, ob die Nullhypothesen das sichere oder das unsichere System unterstellen. Das unsichere System ausschließen zu können ist eine „starke Aussage“ und stellt den positiven Befund dar (Sachs, 1999). Dieser Fall ist das Ergebnis der Verwerfung der einseitigen Nullhypothese, die wie folgt definiert ist:

$$H_0: p > p_0 \tag{2.3}$$

Somit kann aus der Nichtablehnung kein Schluss über die Sicherheit gezogen werden, wenn aufgrund zu geringer Stichprobenzahlen und zu kleiner Teststärke keine Aussage über die Erfüllung einer gegebenen Zielsicherheit getroffen werden kann. Die 4 möglichen Fälle, die durch eine experimentelle Bestimmung auftreten können, sind in Tabelle 3 zusammengestellt.

Tabelle 3: Mögliche Fälle der Entscheidung über die Erfüllung der Zielsicherheit

	$H_0: p > p_0$	$H_1: p \leq p_0$
$\bar{X}_n > p_0$	System korrekt als unsicher erkannt	System fälschlicherweise als unsicher erkannt (falsch negativ)
$\bar{X}_n \leq p_0$	System fälschlicherweise als sicher erkannt (falsch positiv)	System korrekt als sicher erkannt

2.3 Die Abhängigkeit der Irrtumswahrscheinlichkeit zur Anzahl der Stichproben

2.3.1 Die Wahl der Irrtumswahrscheinlichkeit

Die Deklaration eines soziotechnischen Systems als ein sicheres System erfolgt wie bereits erläutert anhand der Erfüllung einer Zielsicherheit. Die Irrtumswahrscheinlichkeit (falsch positiv) beschreibt dabei die Wahrscheinlichkeit für den Auftritt des α -Fehlers:

$$\alpha := P(p > p_0 \wedge \bar{X}_n \leq p_0) \tag{2.4}$$

Die Wahl einer geeigneten maximal zulässigen Irrtumswahrscheinlichkeit ist durch den Anwender von AccSis zu treffen.

2.3.2 Die Schätzung der Abhängigkeit anhand des Tschebyscheff-Ansatzes

Eine erste Abschätzung der Abhängigkeit der Irrtumswahrscheinlichkeit gelingt mit Hilfe dieses Ansatzes, der die Wahrscheinlichkeit für das Übertreten einer maximal zulässigen Toleranzwahrscheinlichkeit ϵ in Abhängigkeit zu der Stichprobenzahl beschreibt. Die Vergrößerung der Stichprobenmenge bewirkt laut dem *Schwachen Gesetz der großen Zahlen* eine Annäherung der gemessenen relativen Häufigkeit \bar{X}_n an die Wahrscheinlichkeit p des Unfallereignisses, beschrieben durch die *Stochastische Konvergenz* nach Jakob Bernoulli, gemäß

$$\lim_{n \rightarrow \infty} P(|\bar{X}_n - p| < \epsilon) = 1, \tag{2.5}$$

liegt der Fehler oberhalb einer tolerierbaren Grenze, kann von dem Effekt der Rechtzensur gesprochen werden. Dieser Zusammenhang lässt sich anhand der *Tschebyscheff-Ungleichung* nachvollziehen, die eine Konvergenz der beobachteten Unfälle X_n an den Erwartungswert μ mit Hilfe eines absoluten Toleranzbereiches a gemäß

$$P(|X_n - \mu| \geq h) \leq \frac{\text{Var}(X_n)}{h^2} \quad (2.6)$$

mit $h \approx n \cdot \epsilon$, $h \in \mathbb{N}$ und $\mu = n \cdot p$, $\mu \in \mathbb{N}$ beschreibt. Der durch diese Gleichung geschätzte Fehler wird als *Tschebyscheff-Risiko* bezeichnet und quantifiziert die Irrtumswahrscheinlichkeit. Durch Einsetzen von $\text{Var}(X_n) = n \cdot p \cdot (1 - p)$, $h = n \cdot \epsilon$ und anschließender Division durch n , lässt sich das *Tschebyscheff-Risiko* für binomialverteilte Zufallszahlen herleiten:

$$P(|\bar{X}_n - p| \geq \epsilon) \leq \frac{p(1-p)}{n \cdot \epsilon^2} \approx \frac{p}{n \cdot \epsilon^2}, \quad (2.7)$$

für $p \ll 1$. Die Zielsicherheit sollte sich in der erlaubten Toleranzwahrscheinlichkeit ϵ wiederfinden lassen. Aufgrund des symmetrischen Charakters der Ungleichung könnte hier z.B. $\epsilon = p_0/2$ zur Anwendung kommen. Jedoch kann aufgrund der Unbekannten p der Rückschluss von einer maximalen Irrtumswahrscheinlichkeit auf die notwendige Anzahl an Stichproben nicht gezogen werden. Bei Gegenüberstellung dieser Ungleichung mit dem soeben eingeführten Sicherheitsargument deutet die Konvergenz bezüglich der Stichprobenanzahl auf einen grundsätzlich antiproportionalen Zusammenhang zwischen α und n hin. Auffällig ist dabei die Konvergenz in Abhängigkeit von p , die bei $p = 0,5$ maximal ist. Systeme, die also eine Unfallwahrscheinlichkeit von 0,5 bieten, weisen die stärkste Konvergenz bei gleicher Stichprobenzahl auf. Die Schätzung des Fehlers auf Basis dieser Ungleichung ist allerdings ungünstig bezüglich sehr kleiner Zielsicherheiten, da die Ungleichung $p/(n \cdot \epsilon^2) > 1$ liefern kann.

2.3.3 Die Schätzung der Abhängigkeit anhand des Binomialansatzes

Eine alternative Herleitung für die Quantifizierung ist die Verwendung der Wahrscheinlichkeitsdichte der Binomialverteilung, die der gedächtnislosen stochastischen Eigenschaft des Unfallereignisses entspricht. Anhand der Gleichung der Irrtumswahrscheinlichkeit in Formel 2.4 können zwei Fälle unterschieden werden, die im Folgenden diskutiert werden.

Die Abbildung 18 (rechts) stellt eine grafische Interpretation der Irrtumswahrscheinlichkeit dar, die die Kenntnis über p für den Fall $p > p_0$ unterstellt. Die Irrtumswahrscheinlichkeit ist dann definiert als die kumulative Wahrscheinlichkeit im Intervall $[0; p_0 \cdot n]$. Aufgrund der gegebenen Kenntnis über p entspricht dieser Fall einer Versuchsanordnung nach Vorbild eines Würfelexperiments und wird hier als der Würfelfall bezeichnet. Aufgrund dieser Kenntnis ist dieser Fall für die Risikoanalyse nicht gültig.

Die Abbildung 18 (links) illustriert den Rückschluss von \bar{X}_n auf α anhand eines linksschiefen Verteilungsbeispiels mit einer beispielhaften Verwerfung der Nullhypothese aufgrund $\bar{X}_n < p_0$. Das Beispiel stellt den praktischen Experimentalfall aufgrund der Kenntnis über \bar{X}_n dar. Die Irrtumswahrscheinlichkeit ist in diesem Fall die Fläche der Wahrscheinlichkeitsdichte im Intervall $[p_0; 1]$. Die Bestimmung der Irrtumswahrscheinlichkeit gelingt bei gegebener Stichprobenzahl n und Messhäufigkeit $X_n \leq p_0 \cdot n$ mit Hilfe der kumulativen Verteilungsfunktion und lässt sich anhand eines Integrals beschreiben:

$$P(p > p_0 | \bar{X}_n \leq p_0) = (n + 1) \int_{p_0}^1 I(p, n, X_n) dp \quad (2.8)$$

mit

$$I(p, n, X_n) = \binom{n}{X_n} p^{X_n} \cdot (1 - p)^{n - X_n} . \quad (2.9)$$

Für die kumulative Verteilungsfunktion muss der Term $n + 1$ ergänzt werden, damit

$$1 = (n + 1) \int_0^1 I(p, n, X_n) dp \quad (2.10)$$

gilt. Eine analytische Lösung für dieses Integral, die Stammfunktion, ist für $X_n > 0$ nicht bekannt. Für $X_n = 0$ kann jedoch die folgende exponentielle Zerfallsfunktion für große Stichproben und kleine Zielsicherheiten genannt werden

$$P(p > p_0 | X_n = 0) = (n + 1) \int_{p_0}^1 (1 - p)^n dp = (1 - p_0)^{n+1} \quad (2.11)$$

$$\xrightarrow{n \gg 1} e^{\ln(1-p_0) \cdot n}$$

$$\xrightarrow{p_0 \ll 1} e^{-p_0 \cdot n}$$

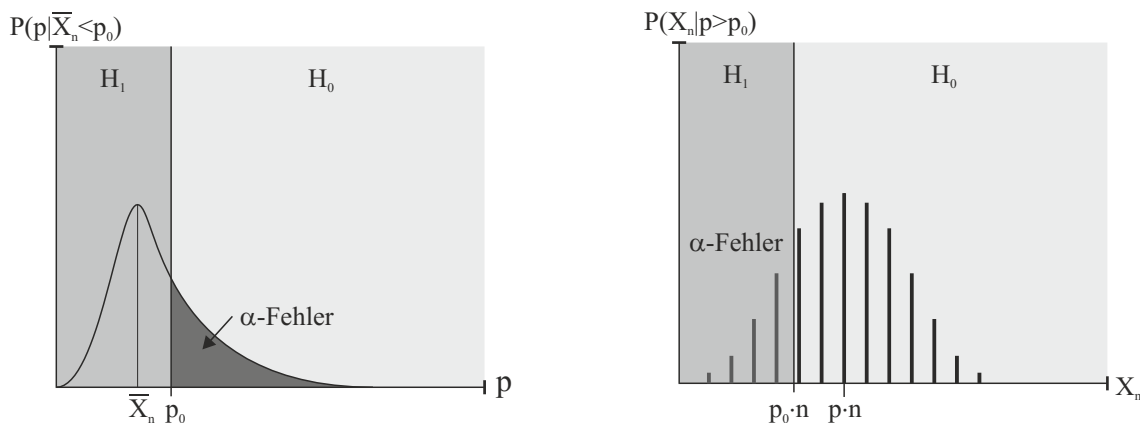


Abbildung 18: Geometrische Interpretation der Irrtumswahrscheinlichkeit des Experimentalfalls (links).
Geometrische Interpretation der Irrtumswahrscheinlichkeit des Würfelfalls (rechts)

Der qualitative Verlauf dieses exponentiellen Zusammenhangs kann als Zerfallsfunktion erachtet werden und weist eine Relaxation mit dem Eigenwert $-p_0$ auf. Dies bedeutet, dass nach $1/p_0$ Stichproben eine Minderung von ca. 63 % erreicht ist. Das Produkt aus $p_0 \cdot n$ bewirkt dabei eine Invarianz der Irrtumswahrscheinlichkeit gegenüber der Wahl der Einheit der Grundgesamtheit, gleich ob es sich hier um Flugstunde oder –minute handelt.

Für eine Abschätzung der Abhängigkeiten der Irrtumswahrscheinlichkeit für den Fall $X_n > 0$ werden zunächst exemplarische numerische Näherungslösungen verwendet, die den quantitativen Verlauf der Konvergenz bei zunehmender Stichprobenzahl illustrieren. Die Abbildung 19, links, veranschaulicht zu diesem Zweck fünf numerische Lösungen der Irrtumswahrscheinlichkeiten, errechnet mit Hilfe der kumulativen Verteilungsfunktion bei

beispielhaften $p_0 = 0,5 \%$ (5:1000) über eine zunehmende Anzahl der Stichproben n . Die gewählte Zielsicherheit wird als hinreichend klein für die Erfüllung der Bedingungen der Formel 2.11 erachtet. Die Wahrscheinlichkeitskurven verdeutlichen das konvergente Verhalten gegen null bei zunehmenden Stichproben. Der unterschiedliche Beginn der Kurven ist durch den jeweiligen Gültigkeitsbereich $X_n \leq p_0 \cdot n$ begründet. Die Zunahme der gemessenen Häufigkeit bewirkt dabei eine Zunahme der Irrtumswahrscheinlichkeit. Anhand der Wahrscheinlichkeitsdichte der Binomialverteilung ist der Trend bei zunehmender Stichprobenanzahl durch eine Verkleinerung der Vertrauensintervalle um \bar{X}_n zu erklären. Die Wahrscheinlichkeit für $P(p > p_0)$ sinkt somit.

Die Ableitung einer dazu passenden analytischen Ersatzfunktion auf Basis der Zerfallsfunktion Formel 2.11 kann zusätzliche Hinweise bieten, welche Gewichtung die Stichprobenanzahl in Relation zur Zielsicherheit und der Anzahl der gemessenen Häufigkeit für die Irrtumswahrscheinlichkeit hat. Die Ersatzfunktion,

$$P(p > p_0 | \bar{X}_n \leq p_0) \approx e^{-p_0 \cdot n + f(X_n)} \quad (2.12)$$

weist als parametrische Asymptotengleichung überwiegende Übereinstimmungen für alle X_n bezüglich des Trends auf. Die Funktion $f(X_n)$ beschreibt dabei eine noch unbekannte Zunahme der Irrtumswahrscheinlichkeit bei steigender gemessener Häufigkeit X_n . Der Vergleich der Zerfallsfunktion mit der numerischen Näherungslösung im XY-Diagramm, Abbildung 19 rechts, verifiziert den kongruenten Trend bei zunehmender Stichprobenzahl für alle X_n . Die numerische Näherungslösung lässt bei zunehmender Stichprobe eine Konvergenz der unbekanntes Funktion $f(X_n)$ gegen einen Grenzwert erkennen. Der Einfluss von X_n kann demnach als systematischer Fehler erachtet werden. Die Zerfallsfunktion liefert trotz der Unbekanntheit von $f(X_n)$ eine Beschreibung für die Konvergenz der Irrtumswahrscheinlichkeit bei steigender Stichprobenzahl. Innerhalb dieser Arbeit wird von einem grundsätzlich exponentiellen Zusammenhang ausgegangen, die die Zielsicherheit als wesentlichen Parameter des Trends herausstellt.

Es sollte allerdings auch berücksichtigt werden, dass die Diskussion der Faktoren auf Grundlage von 5 numerischen Näherungslösungen durchgeführt wurde. Der Rückschluss aus den Näherungslösungen auf die Zerfallsfunktion hat somit keine Allgemeingültigkeit. Allerdings ist die Gegenüberstellung ausreichend, um die exponentielle Konvergenz bei zunehmender Stichprobenzahl bestätigen und die entscheidenden Parameter dieses Trends identifizieren zu können.

Wie bereits in der Einleitung in Kapitel 1 aufgeführt, fordert der Bereich ATM die Erfüllung extrem kleiner Zielsicherheiten von z.B. $1,55e-8$ pro Flugstunde (Eurocontrol, 2001). Der hohe Anspruch an eine experimentelle Beweissicherheit wird am folgenden Rechenbeispiel deutlich: Die Zielsicherheit beschreibt ein Unfallereignis pro 10^9 Stunden. Bei einer gemessenen Häufigkeit des Unfallereignisses von null in 10^9 Stunden beträgt die Irrtumswahrscheinlichkeit bezüglich des α -Fehlers 36,8 %. Um beispielsweise eine 5 % - Wahrscheinlichkeit zu erhalten, werden 5,3-mal so viele Stunden gebraucht.

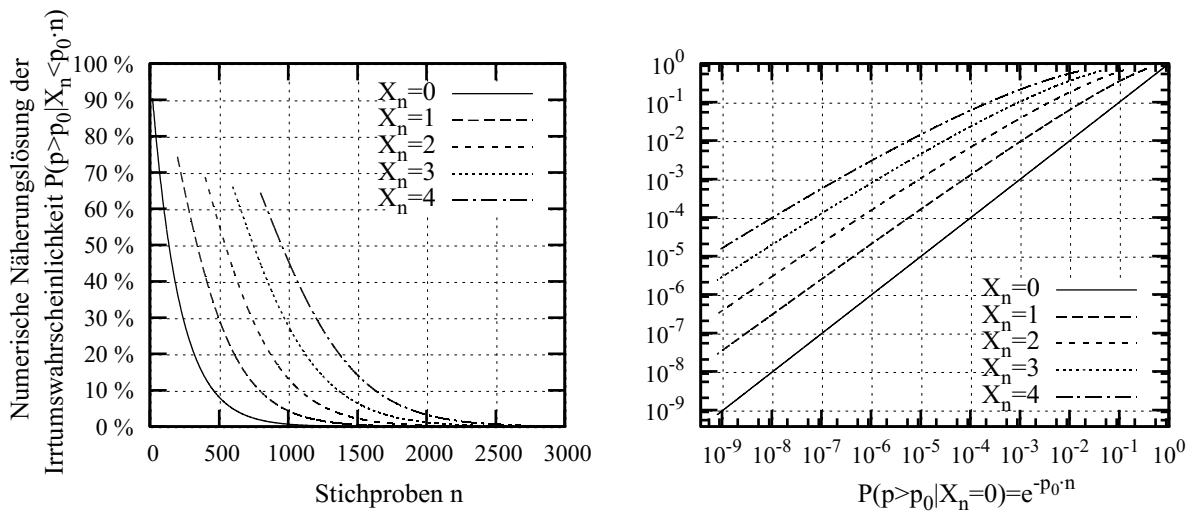


Abbildung 19: Konvergenz der Irrtumswahrscheinlichkeit für fünf hypothetische Messwerte bei einer Zielsicherheit $p_0=0,5\%$ (links). Das XY-Diagramm zeigt den kongruenten Trend der exponentiellen Näherungsfunktion anhand der annähernd linearen Kurven (rechts)

2.4 Ein Konzept zur Beschleunigung der Konvergenz

2.4.1 Die Beschleunigung durch Stresseinwirkung

Die soeben dargelegte Modellierung hat entscheidende Faktoren und Zusammenhänge der Irrtumswahrscheinlichkeit hergeleitet. Den Verfahrensschritten von ALT folgend werden äquivalente Schritte vorgestellt, die die Randbedingungen des soziotechnischen Systems im Flugbetrieb berücksichtigen. Eine offensichtliche Maßnahme zur Minderung der Irrtumswahrscheinlichkeit wäre durch die Erhöhung der Stichprobenzahl n simulierter Stunden zu bewerkstelligen. Eine Alternative zu einer derartigen Maßnahme besteht aus der Beschleunigung der stochastischen Konvergenz. Im Folgenden wird ein Konzept vorgestellt, das sich die Methoden von ALT (Nelson, 2004) zum Vorbild nimmt und eine Methode der Leistungsminderung (*performance degradation*) anwendet.

Die kontrollierte Erhöhung der Unfallwahrscheinlichkeit wird in diesem Konzept mit Hilfe von vordefinierten Stresslasten hervorgerufen, die in ihrer Stärke dosierbar und unterstellt kalibrierbar sind. Da in der Risikoanalyse das Unfallereignis das mögliche Resultat einer Unsicherheit im Betrieb ist, könnten auch alternative Maßnahmen zur Leistungsminderung eingesetzt werden. Im verallgemeinerten Sinne kann man aus diesem Grund bei jedweder Maßnahme von einer Modifikation der Arbeitsrandbedingungen sprechen, die eben diese Unsicherheit kontrolliert fördern. Die daraus entstehende Erhöhung der Unfallwahrscheinlichkeit wird anhand des *Intensivierungsfaktors* a ausgedrückt und ist das Ergebnis der Stresseinwirkung. Der Faktor erhöht die Unfallwahrscheinlichkeit wie folgt:

$$a \cdot p = \hat{p} \quad (2.13)$$

mit $a \leq p^{-1}$. Der Intensivierungsfaktor ist dabei eine unabhängige, extern-wirkende Maßnahme, die während der Versuchsdurchführung die operative Kritikalität durch die Stresseinwirkung erhöht. Bei erfolgter Leistungsminderung bleibt die Hypothesendefinition Formel 2.3 gültig, wenn eine beidseitig symmetrische Skalierung gemäß

$$a \cdot p_0 = \hat{p}_0 \quad (2.14)$$

vorgenommen wird, mit

$$H_0: a \cdot p > a \cdot p_0. \quad (2.15)$$

Ein gezielter Einfluss auf die Geschwindigkeit der Konvergenz wird hier genommen. Ein Faktor für die erhöhte Konvergenzgeschwindigkeit gegenüber Normalbedingungen soll die Beschleunigung ausdrücken: die *relative Intensivierung* γ . Sie ist definiert als der aus der Intensivierung effektiv erzielte Minderungseffekt der Irrtumswahrscheinlichkeit bei gleicher Zahl an Stichproben n , beschrieben durch:

$$\frac{\alpha}{\hat{\alpha}} = \frac{P(p > p_0 \wedge \bar{X}_n \leq p_0)}{P(\hat{p} > \hat{p}_0 \wedge \bar{X}_n \leq \hat{p}_0)} = \gamma \quad (2.16)$$

2.4.2 Die Beschleunigung nach Tschebyscheff

Eine ungefähre Schätzung des zu erwartenden Effekts der Intensivierung liefert die Tschebyscheff-Ungleichung, Formel 2.7, durch Einsetzen der Intensivierung, Formel 2.13, und der skalierten Zielsicherheit, Formel 2.14, gemäß

$$P\left(|\bar{X}_n - p| \geq \frac{\hat{p}_0}{2} = \frac{a \cdot p_0}{2}\right) \leq \frac{\hat{p}}{n \cdot \left(\frac{\hat{p}_0}{2}\right)^2} = \frac{4 \cdot p}{a \cdot n \cdot p_0^2} \quad (2.17)$$

mit der Annahme $\epsilon \approx p_0/2$. Die Intensivierung bewirkt eine Erhöhung der Stichprobenzahl durch $a \cdot n$. Dies lässt einen antiproportionalen Zusammenhang zwischen der Irrtumswahrscheinlichkeit und der Intensivierung vermuten.

2.4.3 Die Beschleunigung nach Binomialverteilung

2.4.3.1 Der Experimentalfall

Eine numerische Verifikation dieses Zusammenhangs liefert Abbildung 20, links, anhand beispielhafter numerischer Lösungen der Irrtumswahrscheinlichkeit für sechs ausgewählte Intensivierungen im Intervall 1 bis 2, berechnet mit Hilfe der kumulativen Binomial-Verteilungsfunktion gemäß Formel 2.8. Die Abbildung illustriert den konvergenten Trend bei zunehmender Intensivierung und Stichprobenanzahl und weist auf das vermutlich exponentielle Verhalten bezüglich a und n hin, wie es in dem vorhergehenden Abschnitt 2.3.3 bereits geschätzt wurde. Zusätzliche Gewissheit bringt die numerische Verifikation im XY-Diagramm in Abbildung 20, rechts, anhand einer Gegenüberstellung der Kurve mit der exponentiellen Zerfallsfunktion

$$P(p > p_0 | \bar{X}_n = 0) \approx e^{-p_0 \cdot a \cdot n}. \quad (2.18)$$

So resultiert ein deckungsgleicher Verlauf aller Kurven mit der numerischen Näherungslösung für alle exemplarisch gesetzten Intensivierungsfaktoren. Die Konsequenz ist ein gerader Verlauf wie in Abbildung 19, rechts, für den Fall $X_n = 0$. Das Ergebnis bestätigt, den Effekt des Intensivierungsfaktors durch die Zerfallsfunktion der Formel 2.18 korrekt zu beschreiben. Eine höhere Unfallwahrscheinlichkeit pro Gelegenheit aufgrund des Intensivierungsfaktors und die Vergrößerung der Stichprobengröße sind demnach effektgleich bezüglich der stochastischen Konvergenz.

Die Unterstellung des Experimentalfalls mit der hypothetischen Messhäufigkeit $X_n = 0$, unterschlägt einen begleitenden Effekt. So bewirkt die Intensivierung durch die erhöhte erwartete Anzahl an Unfallereignissen $\mu = \hat{p} \cdot n$ wiederum eine größere Häufigkeit gemessener Zielereignisse X_n . Dies ist eine praktische Konsequenz der Intensivierung der Unfallwahrscheinlichkeit, die zu einer vermehrten Messung dieses Ereignisses führt. Der Fall $X_n > 0$ ist aber akzeptabel, da die höhere Konvergenz ebenso in diesem Fall vorliegt.

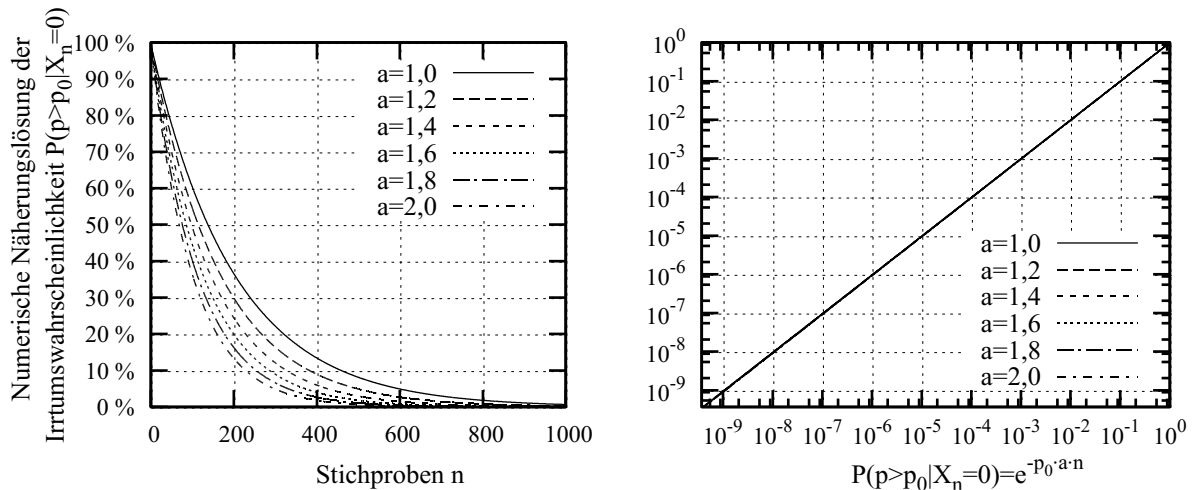


Abbildung 20: Die Irrtumswahrscheinlichkeit für sechs beispielhafte Intensivierungen bei einer Zielsicherheit von $p_0 = 0,5 \%$ und einer hypothetisch gemessenen Häufigkeit $X_n = 0$ (links). Das XY-Diagramm zeigt den getreuen Verlauf der Zerfallsfunktion anhand der annähernd deckungsgleichen Kurven (rechts)

2.5 Ein Konzept zur Stimulation sicherheitsrelevanter Ereignisse

2.5.1 Das Ziel der Stimulation

Der vorhergehende Abschnitt 2.4 hat anhand von Modellentwicklungen gezeigt, dass die Intensivierung der Unfallwahrscheinlichkeit als ein wirkungsvolles Mittel zur Minderung der Irrtumswahrscheinlichkeit und somit des Effekts der Rechtszensur verwendet werden könnte. Als Mittel der Intensivierung wurde einwirkender Stress ausgewählt, der menschliche Fehler gezielt herbeiführen. Der Abschnitt 1.2.1 hat den menschlichen Fehler als eine bekannte und gut erforschte Unfallursache herausgestellt. Ebenso wurde in diesem Abschnitt der wesentliche Beitrag durch Stress hervorgehoben.

2.5.2 Die Wahl des Zeitdrucks als Stressor

Entscheidungen des Menschen sind zahlreichen Stressfaktoren unterworfen, die einen wesentlichen Einfluss auf die Erwägung von Handlungsoptionen haben. Der Begriff „Stress“ entstammt dem lateinischen *stringere*. Die deutsche Sprache bietet mehrere Möglichkeiten der Übersetzungen. Mit dem Hintergrund dieser Arbeit ist eine mögliche passende „drängen“. Eine der ersten Definition von Stress lautete:

„Stress is the non-specific response of the body to any demand made upon it“ (Selye, 1976, S. 137).

Es werden im Folgenden einige Beispiele für arbeitsrelevante Stressfaktoren, sogenannte Stressoren, vorgestellt:

- Unsicherheit und Zeitdruck (Rastegary & Landy, 1993)
- Umgebungsbedingungen, Extrembedingungen und Lebensbedingungen (Green, 1985)
- Stressoren im Bereich der Instandhaltung von Flugzeugen (FAA, 2011)
 - i. Physikalische (Temperatur, Lärm, Licht, enge Arbeitsräume)
 - ii. Psychologische (Finanzielle Probleme, Eheprobleme, soziale Probleme im Arbeitsumfeld)
 - iii. Physiologische (krankheitsbedingte Erschöpfung, Unterernährung, Schlafmangel, unregelmäßige Arbeitsschichten mit z.B. Nachtschichten)

Allgemein kann all jenen Faktoren unserer Umwelt eine Stresswirkung zugesprochen werden, die uns in jedweder Art aktivieren. Aus diesem Grund kann die Menge an Stressfaktoren nicht allgemeingültig eingegrenzt werden.

Stress ist eng verbunden mit dem Konzept der Arbeitsbeanspruchung (*Workload*) und wurde in einer der frühesten Definition als „hypothetisches Konstrukt“ bezeichnet:

„Workload is a hypothetical construct that represents the cost incurred by a human operator to achieve a particular level of performance“ (Hart & Staveland, 1988, S. 2).

Sie ist also ein Maß des mental erfahrenen Stresses bei Durchführung einer spezifischen Aufgabe. Bubb beschreibt demgegenüber Stress als Ursache der Arbeitsbeanspruchung, was dem Konzept der sogenannten Arbeitsbelastung entspricht:

„All external influences on the human are called stress. Depending on the individual characteristics and capabilities of the human, this stress results into individual workload for the respective human operator“ (Bubb, 2005, S. 134)

Unter den zur Auswahl stehenden Stressoren fällt die Wahl auf den Zeitdruck als vielversprechendster Kandidat. Einer der Gründe für die Eignung ist, dass ein simulationsgestütztes Arbeitsumfeld die Induktion von Zeitdruck aufgrund der Kontrolle der Arbeitsbedingungen konzeptionell unterstützt. Zeitdruck liegt innerhalb der Möglichkeiten der Einflussnahme. Ein anderer Grund ist, dass Zeitdruck als Stressor wohl bekannt und gut erforscht ist. Die gegenwärtig bekannten Auswirkungen von Zeitdruck werden zu diesem Zweck im Folgenden kurz vorgestellt.

Hollnagel führt den Zeitdruck als wesentliche Bedingung der Leistungsfähigkeit (*Common Performance Condition*) auf und nennt ihn als wesentliche Ursache für die erzielbare Kontrolle über eine Situation (Hollnagel, 1998). Zeitdruck wird allgemein als ein subjektives Gefühl definiert, weniger Zeit als benötigt zur Verfügung zu haben, um eine Aufgabe zu lösen (Rastegary & Landy, 1993). Eine ähnliche Definition beschreibt Zeitdruck als eine Form von Stress, die auf einen Mangel an Zeit zurückzuführen ist (Freedman & Edwards, 1988). Die menschliche Empfänglichkeit für Zeitdruck ist das Ergebnis einer drohenden Sanktionierung des Menschen bei Überschreiten der verfügbaren Zeit. Somit kann der Zeitdruck als die Wahrnehmung einer nachteiligen Konsequenz erachtet werden. Ist die Entscheidung in einer begrenzten Zeit zu ergründen, muss der Mensch eine Abwägung zwischen der Zeit für die Entscheidungsfindung, die Schnelligkeit, und der möglichen verbundenen Qualität vornehmen. Zeitdruck kann die Qualität der Arbeit und/oder die Leistungsfähigkeit des Operateurs beeinflussen gemäß

$$P = \frac{Q}{t} \quad (2.19)$$

mit der Leistung P , der Qualität Q und der Zeit t (Bubb, 2005). Das Abwägen des Anspruches zwischen *Qualität und Schnelligkeit* ist demnach ein elementares Merkmal jeder Entscheidungsfindung des Menschen, das i.d.R. durch seine Aufgabendefinition im Betrieb einen begrenzten Spielraum zulässt.

Gemäß dieser Theorie ist der Eintritt eines menschlichen Fehlers die Unterschreitung einer minimal geforderten Qualität einer Handlung. Die Wahrscheinlichkeit für den menschlichen Fehler (*Human Error Probability*) wird darin bemessen als

$$HEP = \frac{\text{Anzahl aufgetretener Menschfehler}}{\text{Anzahl der möglichen Gelegenheiten für den Menschfehler}}$$

Dieser Zusammenhang qualifiziert den Zeitdruck als Faktor für die Wahrscheinlichkeit von menschlichen Fehlern. Zeitdruck hat im Allgemeinen einen bedeutenden Einfluss auf das menschliche Verhalten (Edland & Svenson, 1993), wie z.B. die Folgenden:

- Information zu einer Entscheidungssituation werden schneller, dafür aber oberflächlicher zusammengetragen.
- Bei der Gewichtung von positiven und negativen Konsequenzen wird weniger nach Argumenten differenziert und die negativen werden stärker gewichtet.
- Der Mensch schränkt bei Entscheidungsfindungen unter Zeitdruck die verfügbare Lösungsmenge ein und erwägt weniger alternative Handlungsoptionen.

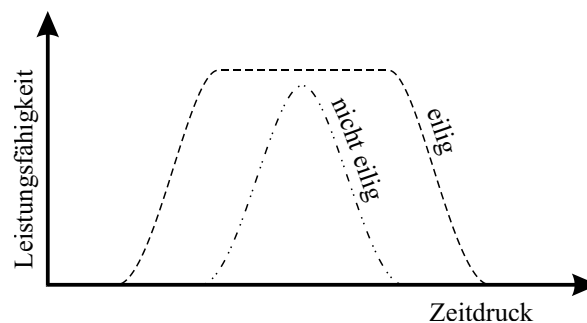


Abbildung 21: Leistungsfähigkeit in Abhängigkeit des vorliegenden Zeitdrucks (Rastegary & Landy, 1993)

Aufbauend auf den Grundlagen des *Yerkes-Dodson-Gesetzes* (Yerkes & Dodson, 1908) konnte eine Leistungskurve, beschrieben durch ein invertiertes U, experimentell nachgewiesen werden (Freedman & Edwards, 1988). Die Abhängigkeit zwischen Leistungsfähigkeit und Zeitdruck kann grundsätzlich in drei Intervalle eingeteilt werden (Abbildung 21):

- *Unterer Lastenbereich*: Zunahme der Leistungsfähigkeit bei niedrigem Zeitdruck
- *Mittlerer Lastenbereich*: konstante Leistung im Mittelintervall
- *Oberer Lastenbereich*: Abnahme der Leistungsfähigkeit bei hohem Zeitdruck

Es wird hierbei unterschieden nach Individualmerkmalen, deren Extreme durch den Typus *time urgent* (eilig) und *nontime urgent* (nicht-eilig) klassifiziert werden (Rastegary & Landy, 1993). *Time-urgent*-Personen zeichnen sich durch eine hohe Lastbasis unzureichender Zeit aus, die einen permanenten Zustand des Zeitdrucks hervorruft. Dieser Typus entwickelt in

Folge eine Robustheit gegenüber variierenden Zeitdrucklasten und bietet eine maximale Leistungsfähigkeit über das Mittelintervall der Kurve. Das nicht-eilige Individuum zeichnet sich demgegenüber durch ein kurzes Intervall maximaler Leistungsfähigkeit aus, bei der Überforderung und Unterforderung ineinander übergehen.

Die Möglichkeit zur Einflussnahme auf die Qualität mit Hilfe einer verfügbaren Zeit qualifiziert den Stressor Zeitdruck für die gezielte Provokation von Menschfehlern. Eine Intensivierung der Unfallwahrscheinlichkeit a erscheint somit möglich.

2.5.3 Der Status Quo der Verfahren zur Induktion von Zeitdruck

2.5.3.1 Die Wechselwirkung zwischen Genauigkeit und Schnelligkeit

Die Induktion von Zeitdruck wurde bereits im Rahmen der Computerinteraktion als Mittel der gezielten Unsicherheitserhöhung erforscht. Dabei konnte die Genauigkeit und die Arbeitsgeschwindigkeit in Relation gesetzt werden, welches unter dem Begriff des *Speed-Accuracy-Tradeoff* (SAT) bekannt ist (Fitts, 1966). Eine mögliche Übersetzung ist die *Abwägung zwischen Genauigkeit und Schnelligkeit*. Der Spielraum des Operateurs, gemäß Formel 2.19 zwischen Qualität und Schnelligkeit abwägen zu können, und die Bedeutung für die zu erwartenden Effekte einer Zeitdruckinduktion, können mit eben solchen SAT-Modellen allgemein erklärt werden. Eine Übertragung der Modelle ist möglich, wenn die Qualität mit der Genauigkeit und die Zeit mit der Schnelligkeit zusammenhängen.

Die meisten Forschungsansätze hatten eine kontinuierliche Funktion eines Stressfaktors als zentralen Untersuchungsgegenstand, wie z.B. den Zeitdruck (Swensson, 1972). Zum Zweck der Beschreibung von zeitdruckabhängigen Entscheidungsprozessen innerhalb sogenannter *Entscheidungsaufgaben* (eng. *choice tasks*) kommen *sequentielle Sample-Modelle* zur Anwendung, die den Verlauf einer Entscheidungswahrscheinlichkeit in Abhängigkeit zur Reaktionszeit darstellen. Sie werden im Folgenden kurz *SAT-Kurve* genannt.

Der SAT kann dabei als das Ergebnis eines äußeren Reizes (*Stimulation*) definiert werden, bei dem der Mensch zur Erfüllung seiner definierten Aufgaben eine bestimmte Zielgenauigkeit in Verbindung mit der dazu gewählten Reaktionszeit erreicht. Die Genauigkeit kann hierbei als ein Maß der Arbeitsqualität erachtet werden. Es zeichnet ein lastenabhängiges Verhältnis zwischen den Ansprüchen nach Genauigkeit und Schnelligkeit aus. Die qualitative Form der Kurve konnte modellbasiert als auch auf Basis empirisch gewonnener Daten (Abbildung 22, Kurve 1) als konkaver Verlauf konstruiert werden (Wickelgren, 1977). Auf empirischer Basis (Abbildung 22, Kurve 2) konnten Kurven konstruiert werden, die einer sigmoidalen Funktion ähneln (Dutilh, Wagenmakers, & Visser, 2011).

Im Gegensatz dazu unterstellt das *Fast Guess Modell* von (Ollman, 1966) eine unstetige Kurve (Abbildung 22, Kurve 4), die aus einem Zweizustands-Wechsel zwischen den beiden Entscheidungsarten *Fast Guess* (schnelle Entscheidungen mit hohem Zufallsanteil der Genauigkeit) und dem *Stimulus Controlled Response* (langsamere Entscheidungen bei hohem Genauigkeitsergebnis) besteht. Ein alternatives *Phase Transition Model* (Dutilh, Wagenmakers, & Visser, 2011) unterstellt eine ähnlich unstetige Kurve (Abbildung 22, Kurve 3). Es ist zu beachten, dass das *Fast Guess Modell* allerdings verworfen wurde (Swensson, 1972), wenngleich die grundlegende Annahme über mögliche Zustandswechsel weiterhin innerhalb der Modelle erhalten blieb.

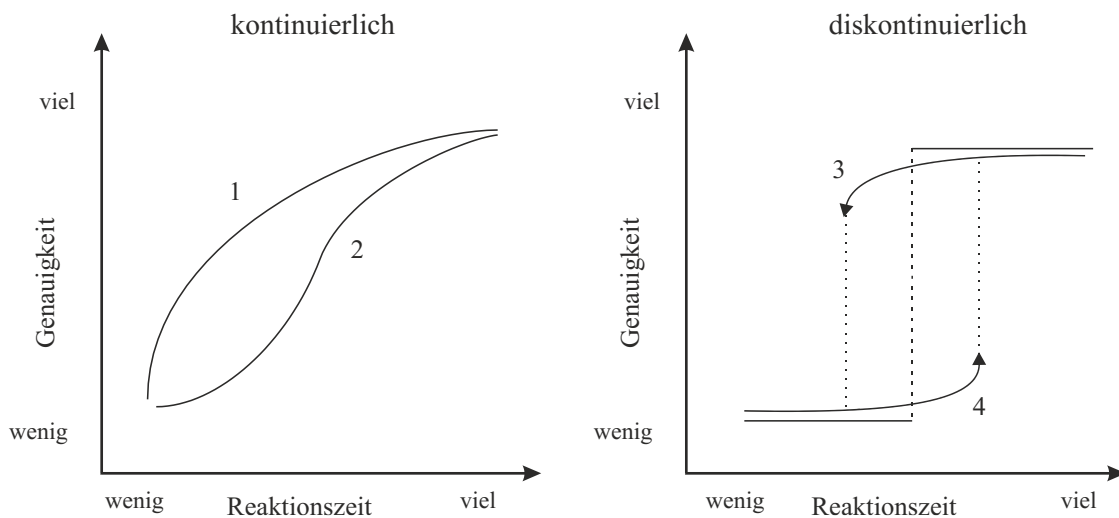


Abbildung 22: Stetige SAT-Kurve, modellbasiertes Verhalten (1 und 2). Unstetige SAT-Kurve (3 und 4) (Dutilh, Wagenmakers, & Visser, 2011)

2.5.3.2 Das Zeitbudget

Die Vorgehensweise zur Induktion von Zeitdruck, die verfügbare Zeit zu verknappen, ist vorgeschlagen in (MacGregor, 1993) und konnte bereits in Studien als methodisches Mittel für die kontrollierte Modifikation der Arbeitsrandbedingungen von Echtzeitsimulationen identifiziert werden, wie z.B. (Svenson & Benson, 1993) und (Bubb & Jastrzebska-Fraczek, 1999). Zentrales Mittel der Einflussnahme ist die gezielte Vorgabe einer maximal verfügbaren Zeit, in der eine Aufgabe zu erfüllen ist. Dies zwingt den Operateur, das Risiko des Zeitfehlers anhand einer Gegenüberstellung von benötigter t_n und verfügbarer Zeit t_a für die Aufgabendurchführung zu schätzen. Als objektives Maß des vorliegenden Zeitdrucks wurde das Zeitbudget definiert (engl. *Time-Constraint Load* oder *Time Budget*), das als Indikation dieses Risikos erachtet wird. Das Zeitbudget setzt die zwei Zeiten in Relation und ist nach (Bubb & Jastrzebska-Fraczek, 1999) wie folgt definiert:

$$TB = \frac{t_n}{t_a} \quad (2.20)$$

Eine Überschreitung der maximal verfügbaren Zeit ist definiert als ein Zeitfehler (*time error*) mit $TB > 1$. Der verwandte Begriff *Time-Constraint Load* hat den Kehrwert des Zeitbudgets zur Definition.

Die Demonstration des starken Zusammenhangs zwischen der verfügbaren Zeit, der benötigten Zeit und dem Zeitbudget zeigen die Ergebnisse einer experimentellen Studie (Bubb & Jastrzebska-Fraczek, 1999). Anhand eines Zeitfensters wurde einer Gruppe von 24 Probanden eine maximal verfügbare Zeit vorgegeben, um Ziffernfolgen mit 3- bis 9-stelligen Nummern auf der Tastatur korrekt abzutippen. Es konnte gezeigt werden, dass die zunehmende Verknappung der verfügbaren Zeit zu einer Beschleunigung der Handlungen der Probanden führte (Abbildung 23, links). Bei Annäherung der beiden Zeiten steigen das Zeitbudget und das Risiko für das Auftreten eines Fehlers bei der Aufgabenausführung (Abbildung 23, rechts).

Für alle Längen der Nummern war auffällig, dass das Zeitbudget bei 50 % einen markanten Punkt darstellt, ab dem die Anzahl der korrekt ausgeführten Handlungen (Leistung) der Probanden einen Einbruch zeigte und die Menschfehlerwahrscheinlichkeit einen Anstieg

erfuhr. Dieser Verlauf deckt sich mit dem Zeitdruck-Leistungs-Zusammenhang nach Rastegary (Abbildung 21) für einen Typus zwischen den Extremen (Rastegary & Landy, 1993).

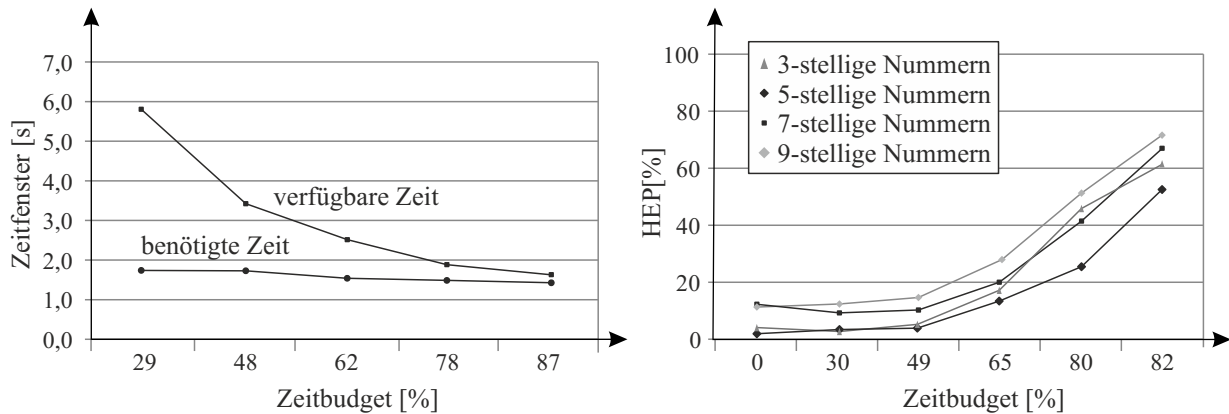


Abbildung 23: Verlauf der benötigten und verfügbaren Zeit über das Zeitbudget für 3-stellige Nummern (links). Verlauf der Menschfehlerwahrscheinlichkeit über das Zeitbudget für 3- bis 9-stellige Nummern (rechts) (Bubb & Jastrzebska-Fraczek, 1999)

2.5.3.3 Die Koexistenz der Induktion mit der Primäraufgabe

Die Aufgabenstellung des Versuchsdesigns wird bei diesem Verfahren vollständig auf die Einhaltung der maximal verfügbaren Zeit ausgerichtet. Sie stellt somit die Primäraufgabe dar und definiert die oberste Priorität, den Zeitfehler zu vermeiden. Beispiele hierfür sind Zeitdruckexperimente, die die Wirkung von Zeitdruck aus unterschiedlichen Motiven untersucht haben, und das Zeitbudgetprinzip erfolgreich zur gesteuerten Induktion verwendeten.

- Testpersonen hatten wie im Vorabschnitt bereits vorgestellt gegebene Ziffernfolgen am Computer innerhalb einer gegebenen Zeit korrekt einzutippen (Bubb & Jastrzebska-Fraczek, 1999).
- Testpersonen hatten Wörter aus einer gegebenen Menge an Buchstaben zu bilden (Freedman & Edwards, 1988). Ziel war die Messung der Erfolgsrate und Leistungsfähigkeit bei variierenden Maximalzeiten.

Weitere Beispiele finden sich in (MacGregor, 1993).

Ein zu implementierendes Verfahren zur Stressinduktion affektiert möglicherweise eine vom Verfahren unabhängig auszuführende Primäraufgabe. Die Einhaltung von verfügbaren maximalen Zeiten als Erweiterung der Primäraufgabe steht im Widerspruch zu den Verfahren, die dem Operateur zur Gewährleistung des konfliktfreien Betriebes in die Hand gelegt wurden, wie z.B. das Holding-Verfahren. Die Implementierung der Zeitdruckinduktion modifiziert die Arbeitsbedingungen des Operateurs wesentlich und stellt vermutlich eine Verzerrung des geltenden Unfallursachenmusters dar. Das ermittelte Risiko bedingt dann diese Modifikation und steht in Konflikt mit dem Ziel, unbedingte Risikowerte zu ermitteln. Eine verträgliche Koexistenz der Primäraufgabe mit einem Zeitdruckinduktionsverfahren scheint aus diesem Grund schwierig zu finden.

2.5.3.4 Die Definition von Zusatzaufgaben

Die Definition einer Aufgabe, die zusätzlich zur Hauptaufgabe zu erfüllen ist, stellt eine Erweiterung des Aufgabenfeldes dar. Sie stellt die Herausforderung für den Operateur, die verfügbare Zeit zwischen den Aufgaben nach Priorität zu teilen (*sequencing*). Die Implementierung in eine Versuchsdefinition wird häufig im Rahmen der Messung menschlicher Leistung für die Indikation der Arbeitsbeanspruchung (*workload*) verwendet (Galster, Duley, Masalonis, & Parasuraman, 2001). Die Leistungsmessung lässt einen Rückschluss auf die verbliebene mentale Kapazität nach Erfüllung der Primäraufgabe zu (Cain, 2007). Die zusätzliche Beanspruchung durch die Zweitaufgabe kann Auswirkungen auf die Hauptaufgaben haben und zu einem unbestimmbaren Messfehler in der ermittelten Arbeitsleistung führen.

Die Definition einer Zweitaufgabe bietet den Vorteil, die Primäraufgabe schützen zu können. Zweitaufgaben, die auch konkrete Handlungen umfassen (z.B. das zusätzliche Bedienen von Knöpfen), haben vermutlich einen wenig kontrollierbaren Einfluss auf die Verknappung der verfügbaren Zeit. So ist sie stark von der menschlichen Fähigkeit abhängig, zu priorisieren und die dazugehörige Aufgabenausführung in den bestehenden Arbeitsfluss aufzunehmen. Dies bestärkt die Anforderung für ein Verfahren, das eine „berührungslose“ Verknappung der verfügbaren Zeit bewirkt und betriebsgebundene Arbeitshandlungen nicht durch Aktivitäten von Zweitaufgaben belastet.

2.5.4 Ein Verfahren zur Induktion von Stress

2.5.4.1 Die Analyse der Verfahren zur Steuerung von Reaktionszeiten

Die zur Ermittlung der SAT-Kurve angewandten Verfahren zur Steuerung der Reaktionszeiten gemäß (Wickelgren, 1977) dienen als Vorbild zur Entwicklung eines geeigneten Induktionsverfahrens. Die folgenden Verfahren werden dabei näher betrachtet:

1. *Anweisungen*: Dies weist den Probanden an, mit einer vorgegebenen Schnelligkeit zu arbeiten (z.B. „So schnell wie es geht!“).
2. *Abfindung*: Dieses Belohnungsprinzip lehnt sich an einem Spielprinzip an, eine auszuzahlende Belohnung in Aussicht zu stellen, die sich bei Lieferung eines „korrekten“ Ergebnisses erhöht und sich im Versagensfall verringert.
3. *Fristsetzung (deadline)*: Die Aufgabenerfüllung muss innerhalb einer gesetzten zeitlichen Frist erfolgen.
4. *Zeitspanne*: Diese Methode erweitert die Fristsetzung um eine Minimalzeit, um eine Mindestqualität zu erzielen.
5. *Antwortsignal*: Diese Methode steuert die entstehende Reaktionszeit anhand von audioakustischen Signalen, z.B. drei Pieptöne in einem vordefinierten Intervall. Nach Vorliegen einer Entscheidungssituation soll die Reaktion des Probanden bei exakt dem dritten Piepton erfolgen. So kann die resultierende Genauigkeit durch Variation der Intervalllängen über ein weites Intervall der Reaktionszeit nachverfolgt werden.
6. *Zerlegung der Reaktionszeiten*: Diese Methode lässt dem Probanden die Freiheit, die Reaktionszeit selbst zu wählen. Die gemessenen Reaktionszeiten werden in vordefinierte Intervallklassen eingeordnet und den gemessenen Genauigkeiten zugeordnet.

Die Verfahren der Punkte 1 bis 5 weisen überwiegende Übereinstimmung mit der Definition einer Zusatzaufgabe entsprechend Abschnitt 2.5.3.4 auf. Sie wären folgerichtig mit der primären Aufgabendefinition einer luftfahrtbezogenen Arbeitsposition vereinbar und Koexistent entsprechend Abschnitt 2.5.3.3. Allerdings kommt bei Punkt 1 zum Tragen, dass eine Reproduzierbarkeit der Reaktionszeiten über längere Zeit vermutlich nicht gelingen bzw. nicht gewährleistet werden kann. Sie lässt dem Menschen einen zu großen Spielraum für die Wahl einer ihm als geeignet erscheinenden Geschwindigkeit. Eine Dosierbarkeit des Zeitdrucks kann somit nicht gelingen. Die Verfahren der Punkte 2 bis 5 unterstellen darüber hinaus die Definition von klar abgrenzbaren Entscheidungssituationen. Die Arbeit einer luftfahrtbezogenen Arbeitsposition unterstützt diese Abgrenzbarkeit nicht. Dieses Merkmal wurde bereits in Abschnitt 2.1.3 beschrieben und als Entscheidungskontinuum bezeichnet. Das Verfahren von Punkt 6 entspricht weitestgehend dem nicht-modifizierten Flugbetrieb, der keine gezielte Intensivierung erfährt.

Zusammenfassend scheint keines dieser Verfahren eine ideale Lösung zu bieten, Zeitdruck während des laufenden Flugbetriebs zu induzieren. Allerdings erscheint das Verfahren unter Punkt 3, Fristsetzung, die größten konzeptionellen Übereinstimmungen mit dem Zeitbudget aufzuweisen. Eine mögliche Weiterentwicklung zur Anpassung an die Problematik des Entscheidungskontinuums würde dieses Verfahren aufgrund der Dosierbarkeit des Zeitdrucks qualifizieren.

2.5.4.2 Ein Induktionsverfahren namens „Konkurrenzdruck“

Eine mögliche Lösung bietet an dieser Stelle die Empfänglichkeit des Menschen für den Wettbewerb. Die Erregung des Wettbewerbes (*competitive arousal*) steht in einem engem Zusammenhang mit dem Gefühl für Zeitdruck (Kerstholt, 1994). Die Rivalität zwischen z.B. Bietern auf einer Auktion und das vorliegende Zeitbudget erhöhen die Wahrscheinlichkeit, wieder zu bieten (Malhotra, 2010). Der empfundene Zeitdruck, Entscheidungen zu treffen, steigt somit in einer Wettbewerbssituation aufgrund des Konkurrenzdrucks. Einen Hinweis auf den starken Zusammenhang zwischen Konkurrenzdruck und Zeitdruck liefert eine Studie, die den Zusammenhang mit Hilfe einer Befragung von 49 Stelleninhabern im Arbeitsleben über gewerblich-industriellen Tätigkeiten mit einer Korrelation von 70% bestimmt (Buch, 2007). Dieser Zusammenhang qualifiziert sich insbesondere aufgrund der Steuerbarkeit der Geschwindigkeit des Wettbewerbs, die den Menschen zu einer zeitkonformen Entscheidungsfindung zwingt.

Die Übertragung auf den luftfahrtbezogenen Anwendungskontext soll mit Hilfe des nun vorgestellten und eigens entwickelten Verfahrens namens Konkurrenzdruck gelingen. Das dazugehörige Akronym ComPerf basiert auf dem englischen Äquivalent *Competitive Performance*, das im weiteren Verlauf der Arbeit zur Anwendung kommt. Arbeiten der Operateur und ein hypothetischer Referenzoperator unter gleichen Anfangs- als auch Randbedingungen, erfüllt dies die Voraussetzung für einen fairen Wettbewerb. Der Vergleich der Leistungsfähigkeiten anhand der Produktivität wird somit möglich. Der Vergleich kann mit Hilfe einer Leistungsmetrik erfolgen, die eine Indikation für die relative Schnelligkeit der Entscheidungsfindung der Operateure untereinander bietet. Die Abbildung 24 zeigt diesen Wirkungsmechanismus des Wettbewerbs, bei der sich der menschliche Operateur und der Referenzoperator gegenüberstehen.

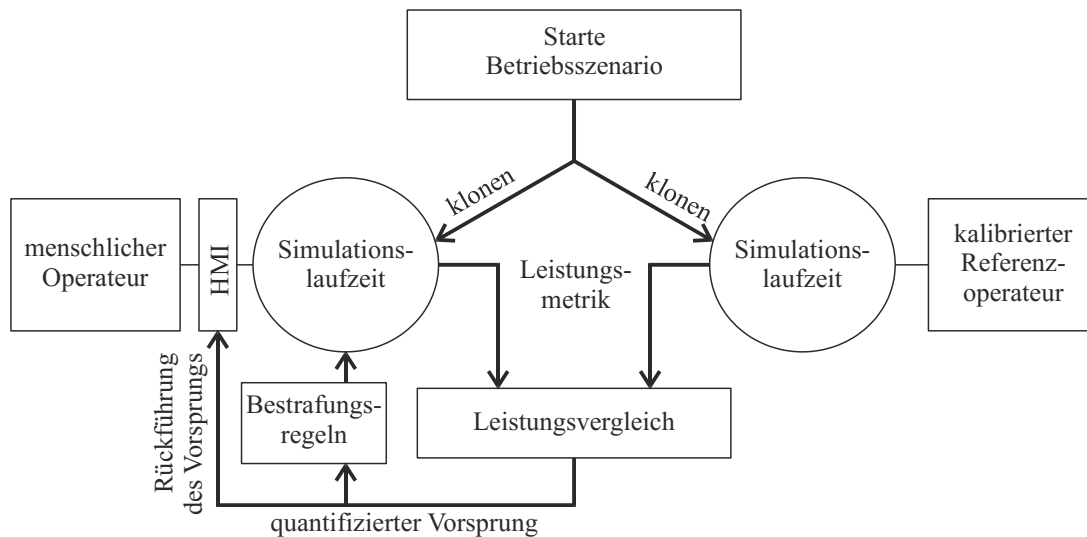


Abbildung 24: Das Wettbewerbsverfahren setzt den menschlichen Operateur und einen Referenzoperateur mit Hilfe eines Leistungsvergleiches in eine Wettbewerbssituation

Der durch den Konkurrenzdruck aufgebaute Zeitdruck wird mit Hilfe einer definierbaren Arbeitsgeschwindigkeit des Referenzoperateurs gesteuert, die als unabhängige Größe zu setzen ist. Die Arbeitsgeschwindigkeit ist dabei definiert als eine Reaktionszeit pro Entscheidung. Wird die Reaktionszeit des Referenzoperateurs in Verhältnis mit der menschlichen Arbeitsgeschwindigkeit gesetzt, so kann dieses als Zeitbudget mit der Reaktionszeit des Referenzoperateurs als verfügbare Zeit t_a erachtet werden. Das Setzen dieser verfügbaren Zeit steuert durch das damit verbundene Zeitbudget die geltenden Randbedingungen des simulierten Betriebs aus einem „Normal“ aus. Die verfügbare Zeit wird aus diesem Grund als „Zeitdrucklast“ bezeichnet, da sie eine Bedingung für die Arbeit des Menschen ist, die belastet. Bei Verwendung des Induktionsverfahrens ersetzt der Begriff den allgemeineren Begriff der „Stresslast“, wie er stets zuvor benutzt wurde.

2.5.4.3 Die Wechselwirkung zwischen der Basislast und dem induzierten Zeitdruck

Eine betriebsbedingte Basislast des Operateurs entsteht grundsätzlich durch die Ausübung einer Aufgabe. Diese wird als eine aufgabenbedingte Arbeitsbelastung erachtet, deren Ausmaß den Designstress definiert. Welche damit verbundene subjektive Arbeitsbeanspruchung der Operateur während des Regelbetriebes maximal erfährt, wäre durch eine Arbeitsbeanspruchungsanalyse nach Vorbild von z.B. (Eurocontrol, 2005) zu ermitteln. Die Auslegung der objektiven Grundlast des Operateurs ist im hohen Maße vom Betriebskonzept des soziotechnischen Systems abhängig und bedingt ebenso äußere Wetter- bzw. betrieblichen Störeinflüsse. Die Deutsche Flugsicherung GmbH legt beispielsweise die Sektorengrößen der Streckenlotsen und die damit verbundenen Kontrollhandlungen pro Zeit derart aus, dass ein Zeitbudget von maximal 70 % erreicht wird (Platz, 1991). Das Zeitbudget ist dabei üblicherweise als Zeit mit messbaren Handlungen des Operateurs pro Stunde definiert. Zu einem gleichwertigen Ergebnis kommt auch die Konzeptstudie von Eurocontrol, die eine Belastung oberhalb von 70 % als Überbelastung erachtet (Eurocontrol, 2005). Ein Zeitbudget

von 70 % kann demnach als maximale Obergrenze der zulässigen Arbeitsbelastung von Streckenlotsen erachtet werden.

Bei Vorgabe einer abnehmenden verfügbaren Zeit durch das Zeitdruckinduktionsverfahren entsteht vermutlich eine Wechselwirkung zwischen einer solchen „Basislast“ und dem vermittelten Zeitdruck, wobei der induzierte Zeitdruck der Basislast zu einem unbekanntem Zeitbudget TB überlagert wird. Im hohen Lastbereichen wird der induzierte Zeitdruck aufgrund der zu erwartenden höheren Stressreaktion vermutlich als dominierend erachtet. Im unteren Lastbereich jedoch ist zu vermuten, dass der effektiv vermittelte Zeitdruck bei zunehmender verfügbarer Zeit t_a durch die Basislast als unteren Grenzwert beschränkt wird (Abbildung 25, rechts). Diesen Bereich kennzeichnet einfach gesagt das Empfinden des Operators, der Zeitdrucklast durch das Induktionsverfahren aufgrund der hohen verfügbaren Zeit keine Bedeutung schenken zu müssen. Dieser Grenzwert entsteht dabei gemäß

$$\lim_{t_a \rightarrow \infty} TB = \frac{t_{n0}}{t_{a0}}. \quad (2.21)$$

Die unbekannte Wechselwirkung führt für das Konzept zu zwei Problemen:

- Es könnte sich hier um einen unstetigen oder um einen gleitenden Übergang handeln. Abbildung 25, rechts, unterstellt einen übergangslosen Punkt zwischen den beiden Zeitdrucklasten. Dem Operator ist es dabei überlassen, welche verfügbare Zeit er effektiv verspürt und als ausschlaggebend für seine Handlungen erachtet. Dies kann sehr individuell ausfallen.
- Anhand der verfügbaren Zeit des Referenzoperators ist nicht nachvollziehbar, welcher Zeitdruck nun effektiv vermittelt wird. Der induzierte Zeitdruck ist somit durch die bloße Festsetzung einer verfügbaren Zeit nicht zwangsläufig auf die Basislast kalibriert.

Ein nützliches Bindeglied der beiden Zeitdrucklasten ist die benötigte Zeit. Wie die experimentelle Studie nach Bubb (Abbildung 25, links) gezeigt hat, ist die Abnahme der benötigten Zeit eine Reaktion auf die abnehmende verfügbare Zeit. Dies kann ebenso für die Basislast t_{a0} und die dazugehörige mittlere benötigte Zeit t_{n0} gelten, deren Quotient das Zeitbudget der aufgabenbedingten Arbeitsbelastung und somit den Grenzwert durch die Basislast bildet (Abbildung 25, rechts). Die benötigte Zeit resultiert demnach aus der empfundenen verfügbaren Zeit und indiziert den effektiv vermittelten Zeitdruck. Aufgrund dieses markanten Punktes stellen die Parameter t_{a0} und t_{n0} charakteristische Grenzwertgrößen des Designstresses dar. Die Bestimmung insbesondere des Parameters t_{a0} ist entscheidend zur Kalibrierung der induzierten Zeitdrucklast gegenüber dem Designstress.

2.5.4.4 Die Bestrafung

Fällt der Leistungsvergleich zu Ungunsten des menschlichen Operators aus, wird dies als Verletzung der verfügbaren Zeit gewertet. Dieses Ereignis bildet den Zeitfehler. Die Sanktionierung erfolgt mit Hilfe einer Bestrafung, die dem Betrieb des menschlichen Operators zugeführt wird. Diese Bestrafung sollte die potentiellen Ursachen eines Unfalls

nicht berühren und berührungslos für den Betrieb wirken. Dies kann zum Beispiel durch die zeitliche Verlängerung des simulierten Arbeitseinsatzes gewährleistet werden.

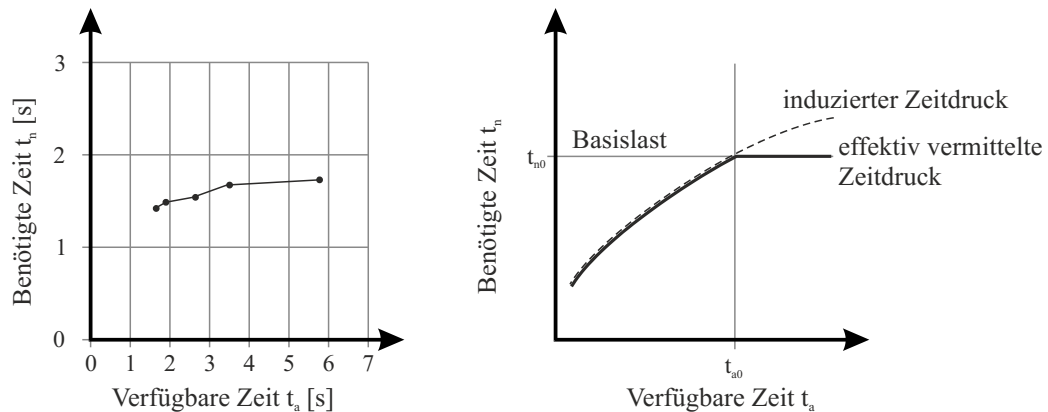


Abbildung 25: Benötigte Zeit als Stressreaktion der verfügbaren Zeit (Bubb & Jastrzebska-Fraczek, 1999) (links). Wechselwirkung des induzierten Zeitdrucks mit der Basislast im Übergangsbereich (rechts)

2.6 Die Regressionsanalyse zur Verifikation der Zielsicherheit

Dieser Abschnitt stellt grundlegende Theorien und Modelle vor, die für eine Verifikation der Zielsicherheit notwendig ist.

2.6.1 Die Regressionsanalyse zur Bestimmung der Unfallereignisverteilung

Könnte eine Anzahl von Unfällen beobachtet werden, ist eine parametrische Beschreibung der Verteilung zu der jeweils vorliegenden Zeitdrucklast möglich. Die Methodik ATL verweist zu diesem Zweck auf die Regression anhand eines Verteilungsmodells, das die Sensitivität des Intensivierungsfaktors zu einer gegebenen Zeitdrucklast berücksichtigt. ALT empfiehlt die Verwendung parametrischer oder nicht-parametrischer Verteilungen zur Nachbildung der *wear-out*-Kurve, die den Ausgangspunkt für den Rückschluss aus den aufgezeichneten Kurven auf die Bedingungen des Designstresses bildet. Die für die Regressionsanalyse unterstellte Datenbasis ist

- eine Anzahl von Eintrittshäufigkeiten von Unfällen X_n ,
- als Ergebnis von einer Anzahl unabhängig durchgeführter Simulationen
- mit je einer Durchführungslänge t_{sim} und
- unter Bedingungen einer konstanten Zeitdrucklast t_a .

Das Verteilungsmodell bildet die Häufigkeitsverteilung von X_n nach, deren Form in einem bestimmten Zusammenhang zur Zeitdrucklast t_a steht (vgl. Abbildung 26).

Im Folgenden werden zwei Verteilungen zur Verwendung als Regressionsmodell diskutiert:

- *Binomialverteilung*: Die Verwendung einer parametrischen Verteilungsbeschreibung beruht auf der Annahme, aufgezeichnete Unfallereignisse als eine Aneinanderreihung unabhängiger Bernoulli-Verteilung zu beschreiben. Diese Annahme wurde bereits in Abschnitt 2.1.4 vorgestellt. Die Schätzung des Parameters p zur Beschreibung der Verteilungsform könnte dabei wie folgt erfolgen:

$$p \approx \bar{X}_n. \quad (2.22)$$

Bei der Schätzung von n zeigt sich ein Problem der Definition der Grundgesamtheit, beschrieben in Abschnitt 2.1.3. Die Schätzung der Stichprobenrate n ist variant zur gesetzten Zeiteinheit, hier in Flugstunden angenommen. Ebenso führt die Dreidimensionalität der Binomialverteilung der drei Faktoren p , n , und X_n zu einer Überparametrisierung, wenn die Datengrundlage lediglich eine eindimensionale Verteilung beschreibt. Eine Parametrierung der Binomialverteilung ist demnach nur möglich, wenn die statistische Einheit für die Gelegenheit eines Unfalls klar abgrenzbar beschrieben ist und die Grundgesamtheit somit keinem Entscheidungskontinuum unterliegt. Falls diese Voraussetzung nicht erfüllt ist, wird die Beschreibung der entstehenden Verteilung von Unfallereignissen mit Hilfe der Binomialverteilung nicht empfohlen.

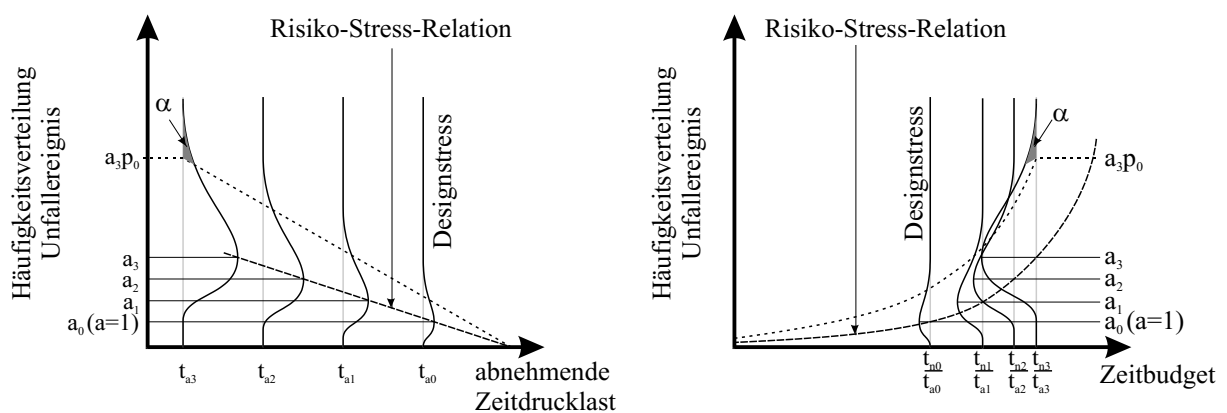


Abbildung 26: Konzept der stressinduzierten Intensivierung der Wahrscheinlichkeit des Unfallereignisses über drei hypothetische Zeitdrucklaststufen (links) und über das Zeitbudget (rechts)

- *Poissonverteilung*: Die Erfassung der Häufigkeit einer Anzahl an Simulationsläufen bildet die Basis für diese stetige parametrische Verteilung, die den Grenzwert einer unendlichen Anzahl an Bernoulli-Experimenten darstellt. Die gültige Anwendung lässt sich mit Hilfe des *Gesetzes der kleinen Zahlen (Law of Raw Events)* erklären:

„The law of rare events states that the total number of events will follow, approximately, the poisson distribution if an event may occur in any of a large number of trials but the probability of occurrence in any given trial is small.“ (Cameron, 2013, S. 5)

Die parametrische Modellierung durch Poissonverteilungen konnte anhand der indirekten Risikometriken Runway Incursion und Staffelungsunterschreitung in (Meyer, Schultz, & Fricke, 2011) nachvollzogen werden und bestätigt die gedächtnislose Stochastik von sicherheitsrelevanten Ereignissen anhand der Relation zur Exponentialverteilung:

„if the poisson provides an appropriate description of the number of occurrences per interval of time, then the exponential will provide a description of the length of time between occurrences.“ (Cooper, 2005, S. 123).

Die Schätzung der Verteilungsform der Poissonverteilung auf Basis eines Parameters λ wird unterstützt durch die eindimensionale Verteilung der Ereignisdaten. Die Beschreibung entspricht dann

$$f(X_n) = \begin{cases} \frac{\lambda^{X_n} \cdot e^{-\lambda}}{X_n!} & \text{für } X_n = 1, 2, 3, \dots \\ 0 & \text{sonst} \end{cases} \quad (2.23)$$

- Alternativ können nicht-parametrische Verteilungsformen gewählt werden, die kein A-priori-Wissen über eine Form der Verteilung verlangen. Dieses kann z.B. anhand des Medians und der Quantile der Verteilung erfolgen. Die Werte würden äquivalent zu einer parametrischen Verteilungsbeschreibung behandelt.

Die resultierende Anpassungsgüte wäre anhand eines statistischen Tests wie z.B. dem Chi-Quadrat-Test zu prüfen und wäre zur Wahl eines geeigneten Verteilungsmodells ausschlaggebend, wenn neben der Poissonverteilung alternative Verteilungsmodelle in Erwägung gezogen werden. Bei einer größeren Anzahl an Simulationsläufen und kleinen Wahrscheinlichkeiten bietet sich aufgrund der Grenzwertregel die Poissonverteilung an.

2.6.2 Die Regressionsanalyse zur Verifikation der Zielsicherheit

Zum Zweck des Rückschlusses der beobachteten Unfallereignisverteilung auf die Verteilung zu Bedingungen des nichtmodifizierten Betriebs (Designstress) soll eine Regressionsanalyse zur Anwendung kommen. Die Basislast als Bezugsgröße des Designstresses wurde in Abschnitt 2.5.4.1 vorgestellt. Es wird an dieser Stelle unterstellt, dass ein Kalibrierungsverfahren die Grenzwertgrößen t_{a0} und t_{n0} der Basislast liefern kann. Alle angefahrenen Zeitdrucklasten können somit in Bezug zu diesen Größen gesetzt werden. Die für die Regressionsanalyse verwendeten Stützstellen gelten somit als kalibriert. Die allgemeine Abhängigkeit zwischen der Zeitdrucklast und der Unfallwahrscheinlichkeit im oberen Lastenbereich soll dabei anhand eines Regressionsmodells beschrieben werden: der *Stress-Risiko-Relation*. Die Abbildung 26, links, illustriert das Prinzip der Regressionsanalyse anhand eines exemplarischen linearen Regressionsmodells.

Je nach physikalischen Gesetzen und Stressverfahren liefern Studien zu beschleunigten Stresstests eine breite Auswahl an Modellen, denen im jeweiligen Anwendungsfeld Gültigkeit unterstellt wird oder aber auch teilweise empirisch nachgewiesen sind, wie z.B. bei ALT die Temperatur-Alterungs-Kennlinie (*Arrhenius Gleichung*) oder das halb-parametrische *Proportional Hazards Model* nach Cox (Cox, 1972). Bei einfachen Relationen können lineare oder exponentiell-lineare Funktionen eingesetzt werden.

Das konzeptionelle Merkmal einer geeigneten Stress-Risiko-Relation ist die Beschreibung eines oder mehrerer Parameter der Unfallverteilung in Abhängigkeit einer gesetzten Zeitdrucklast t_a . Diesen systematischen Zusammenhang stellt eine Verteilungsfunktion $f(t_a)$ her, die die Steigerung der Unfallwahrscheinlichkeit durch die Intensivierung berücksichtigt. Bisher ist kein derartiger Zusammenhang systematisch untersucht und in ein Regressionsmodell überführt worden. Dies mag daran liegen, dass der Zusammenhang zwischen menschlichen Fehlern und Unfallwahrscheinlichkeit hinsichtlich des Einflusses der Intensität des wirkenden Stresses wenig erforscht worden ist.

Im Folgenden werden zwei Risikomodelle vorgestellt, die die kausale Verkettung zwischen Zeitdruck und Unfallwahrscheinlichkeit sowie Beschleunigung abbilden.

2.6.2.1 Modellierung des einfachen Risikomodells

Die Verkettung der aufgeführten Wirkmechanismen, die vom induzierten Zeitdruck bis zur Intensivierung der Unfallwahrscheinlichkeit und zu der Beschleunigung reicht, bildet ein einfaches Risikomodell (Abbildung 27). Die Verkettung besteht im Wesentlichen aus den drei folgenden Schritten:

1. *Zeitdruckinduktion*: Setzen der verfügbaren Zeit und Herstellen eines Zeitbudgets
2. *Fehlerstimulation*: Herabsetzen der Leistungsfähigkeit und der Qualität menschlichen Handels
3. *Intensivierung und Beschleunigung*: Intensivierung der Unfallwahrscheinlichkeit, Minderung der Irrtumswahrscheinlichkeit und somit Beschleunigung des Tests

Die Wirkungskette bildet somit den elementaren Zusammenhang für die Realisierung des Stresstestkonzepts von AccSis.

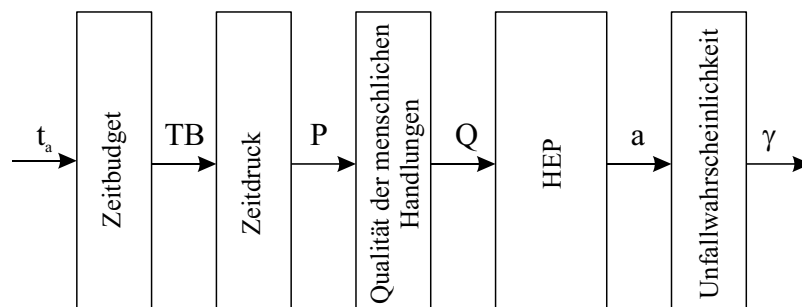


Abbildung 27: Einfaches Risikomodell

Die Verkettung der Einzelmechanismen unterstellt eine Eindimensionalität der Wirkungsrichtung, die die Vielfältigkeit der möglichen Konsequenzen und Quereinflüsse von Randbedingungen nicht berücksichtigt. Mögliche Randbedingungen für die Stärke der ausfallenden Intensivierung können z.B. aus einem variierenden Individualverhalten resultieren, den Zeitdruck auch bei Herstellung gleicher Arbeitsbedingungen nicht reproduzieren zu können. Ebenso sind Störeffekte durch Kompensationsmechanismen des Menschen vorstellbar, dem Zeitdruck mit überlegten bzw. angelegten Strategien zu begegnen und auszugleichen. Welchen tatsächlichen Effekt eine variierende verfügbare Zeit auf die Unfallwahrscheinlichkeit hat, kann durch die bislang bekannten Ergebnisse der Grundlagenforschung nicht vorhergesagt werden.

2.6.2.2 Modellierung des erweiterten Risikomodells

Die Abwägung zwischen Genauigkeit und Schnelligkeit durch den Operateur und die Relevanz für das vorgestellte Zeitdruckinduktionsverfahren wurden bereits in Abschnitt 2.5.3.1 vorgestellt. Sie wird als relevanter Einfluss auf die Wirkungskette zur Intensivierung der Unfallwahrscheinlichkeit erachtet. Demnach ist zu vermuten, dass diese beiden Größen in einem starken Zusammenhang betrachtet werden müssen, um die Wirkung des Zeitdrucks auf

den Menschen und die resultierende Intensivierung der Unfallwahrscheinlichkeit zu verstehen und zu begründen. Eine Erweiterung des einfachen Risikomodells soll die Abwägung aus diesem Grund mit berücksichtigen.

Die Möglichkeit des Operators, zwischen Qualität und Schnelligkeit abzuwägen, eröffnet einen zweiten Wirkungspfad, die benötigte Zeit durch rasches und zügiges Handeln zu verkleinern und die Produktivität somit zu steigern (Abbildung 28). Die benötigte Zeit führt demnach zu einer Produktionsrate mit dem Produktionsgut n pro benötigte Zeit t_n . Es handelt sich dabei um eine Produktion, die entsprechend den Anforderungen des jeweiligen Aufgabenkontextes festzulegen ist, wie z.B. kontrollierte Verkehrsbewegungen pro Zeit oder die Abarbeitung von Arbeitsschritten pro Zeit. Die Rückkopplung der benötigten Zeit begründet das Zeitbudget als Referenzgröße der Zeitdruckinduktion. An dieser Stelle zeigt sich der konzeptionelle Vorteil dieses Modells, das die benötigte Zeit als Teil der Wirkungskette erfasst und der Wirkungskreislauf anhand des Zeitbudgets geschlossen wird. Die unzulässige Überschreitung der verfügbaren Zeit definiert schließlich den Zeitfehler als dritte Ausgangsgröße.

Aufgrund der Kopplung zwischen der entstehenden Unfallwahrscheinlichkeit und der Schnelligkeit, kann auf Basis dieses erweiterten Risikomodells von einer Abwägung zwischen der *Risikobereitschaft und Schnelligkeit* gesprochen werden. Der Begriff der „Bereitschaft“ bringt den individuellen Freiheitsgrad des Operators zur Geltung, durch eine Verringerung der Arbeitsqualität schnellere aber auch riskantere Entscheidungen treffen zu können.

2.6.2.3 Entwicklung eines Regressionsmodells

Welchen qualitativen Verlauf ein Zusammenhang auf Basis der soeben vorgestellten Risikomodelle nehmen könnte, lässt sich anhand des Zusammenhangs zwischen Leistung und Zeitdruck nach Rastegary abschätzen (Rastegary & Landy, 1993). Dieser Zusammenhang wurde bereits im vorhergehenden Abschnitt 2.5.2 anhand der invertierten U-Kurve erläutert (Abbildung 21). Der allgemeine Zusammenhang zwischen Qualität, Leistung und Zeit nach Bubb (Formel 2.19) liefert eine wichtige kausale Verknüpfung zum menschlichen Fehler, der als das Ergebnis mangelnder Qualität beschrieben wird. Die Qualität und die Menschfehlerwahrscheinlichkeit korrelieren demnach negativ. Erachtet man den professionellen, geschulten und trainierten Operateur als ein Individuum, das überwiegende Übereinstimmung mit dem „eiligen“ Typus hat, so könnte die Menschfehlerwahrscheinlichkeit einen qualitativen-zeitdruckabhängigen Verlauf wie in Abbildung 29, links, nehmen. Das Zeitbudget t_{n0}/t_{a0} für eine optimale Arbeitsbelastung im Regelbetrieb (Basislast) sollte aus diesem Grund im Mittelintervall liegen, in dem die Menschfehlerwahrscheinlichkeit minimal ist.

Die Überführung von einer Menschfehlerwahrscheinlichkeit zu dem damit verbundenen Risiko kann mit Hilfe der folgenden Annahmen gelingen:

1. Steigender Zeitdruck und steigende Arbeitsbelastung sind über das Zeitbudget positiv linear korrelierend (der Zeitdruck wird demnach als eine Folge der Arbeitsbelastung erachtet).

2. Steigende Arbeitsbelastung führt aufgrund der damit verbundenen steigenden Menge sicherheitsrelevanter Handlungssituationen zu einer linearen Erhöhung der Anzahl an Gelegenheit für den Eintritt des Unfalls pro Zeit.
3. Die Gelegenheit für einen Unfall und einen menschlichen Fehler wird zu jeder sicherheitsrelevanten Handlungssituation unterstellt.
4. Die Unfallwahrscheinlichkeit korreliert positiv mit der Menschfehlerwahrscheinlichkeit.

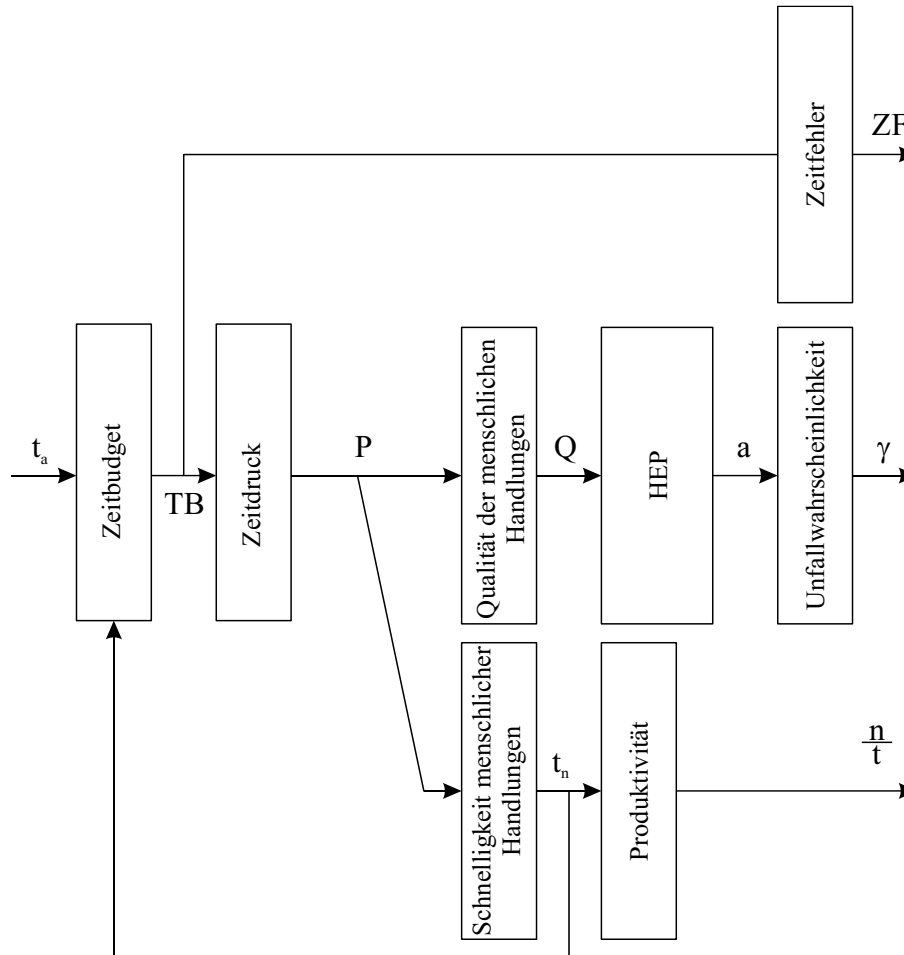


Abbildung 28: Erweitertes Risikomodell

Das Ergebnis der Zusammenführung aller Annahmen wird in Abbildung 29, rechts, dargestellt und dient als vorläufige Ergründung einer Stress-Risiko-Relation. Durch die zunehmende Anzahl an Gelegenheiten pro Zeit bei steigendem Zeitdruck wird der Verlauf der Menschfehlerwahrscheinlichkeit zunehmend stärker gewichtet. Das Modell liefert den Hinweis, dass das Risiko im oberen Lastenbereich in hohem Maße zeitdrucksensitiv zu erwarten wäre. Dieser Bereich bietet demnach ideale Bedingungen für die Steuerung der Beschleunigung und das Anfahren von Zeitdrucklasten.

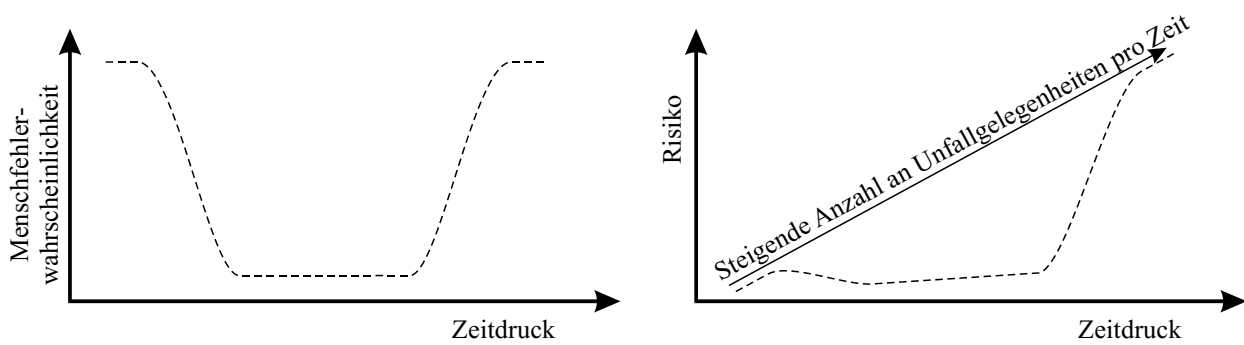


Abbildung 29: Menschfehlerwahrscheinlichkeit abhängig vom Zeitdruck (links). Risiko abhängig vom Zeitdruck (rechts)

Der Verlauf weist eine Übereinstimmung mit dem Untersuchungsergebnis einer Studie auf, bei der die Anzahl der Runway Incursion eine positive lineare Korrelation zu der vorliegenden Verkehrslast aufweist (ALPA, 2007). In dem Informationsbericht wird dieser Zusammenhang mit der zunehmenden Anzahl an möglichen Konfliktszenarien bei zunehmender Anzahl an Verkehrsbewegungen begründet. Diese Begründung erscheint anhand der Anzahl der möglichen logischen Kombinationen plausibel und liefert einen ersten Hinweis auf die mögliche Gültigkeit der vorgeschlagenen Stress-Risiko-Relation.

Allerdings ist an dieser Stelle zu berücksichtigen, dass die soeben diskutierte und vorgeschlagene Relation eine effektive Erhöhung der Arbeitsbelastung (Annahme 1) und somit eine Erhöhung der Gelegenheiten für den Unfall unterstellt (Annahme 2). Diese Annahmen sind kontrovers, da das Zeitdruckinduktionsverfahren prinzipiell keine solchen Gelegenheiten zusätzlich schafft. An dieser Stelle liefert das Grundprinzip der Beschleunigung im Abschnitt 2.4 eine Erklärung. Sowohl die Irrtumswahrscheinlichkeit nach Tschebyscheff (Formel 2.17) als auch nach Herleitung auf Basis der Binomialverteilung (Formel 2.18) weisen die Intensivierung a und die Stichprobenzahl n (hier als Anzahl der Gelegenheiten für den Unfall ausgelegt) als wirkungsäquivalente Faktoren aus. Grundvoraussetzung hierfür ist die Basislast mit einer Arbeitsbelastung zum Regelbetrieb. Es wurde bereits in dem betreffenden Abschnitt 2.4.3.1 geschlussfolgert, dass eine höhere Unfallwahrscheinlichkeit pro Gelegenheit und die kumulative Wahrscheinlichkeit mehrerer Gelegenheiten für den Unfall zu demselben Risiko führen können. Es kann demnach ein ähnlicher Verlauf im oberen Lastenbereich bei induziertem Zeitdruck erwartet werden. Für das gezielte Abtasten dieses Zielbereiches wären geeignete Zeitdrucklasten anzufahren, die als Stützpunkte der Regressionsanalyse dienen. Zu diesem Zweck sollte ein Kalibrierungsverfahren zur Anwendung kommen, das

1. den Zielbereich identifizieren und eingrenzen kann und
2. die Grenzwertgrößen der Basislast t_{a0} und t_{n0} quantifiziert, um den induzierten Zeitdruck gegenüber der Basislast zu kalibrieren.

Für einen guten Kompromiss zwischen Aufwand bzw. Kosten und dem entstehenden Nutzen werden 2 bis 3 Stützstellen in dem Zielbereich empfohlen (Nelson, 2004).

2.6.2.4 Verifikation der Zielsicherheit

Die Regressionsanalyse mit Hilfe der Stress-Risiko-Relation liefert Parameter zur Beschreibung der Unfallverteilung zu Bedingungen des Designstress. Anhand des

Funktionswertes von p_0 der komplementären kumulativen Verteilung $F(p_0) = 1 - P(p \leq p_0)$ kann die Irrtumswahrscheinlichkeit quantifiziert werden, die bei Annahme des sicheren Systems gilt (falsch positiv). Abbildung 26 gibt zu diesem Wert an der Markierung α eine grafische Interpretation. Die Wahl eines geeigneten Vertrauenswertes ist den Anforderungen des jeweiligen Anwendungskontexts anzupassen.

2.7 Die Verifikationsziele

Die Risikomodelle des Abschnitts 2.6.2 und das Zeitdruckinduktionsverfahren „Konkurrenzdruck“ sind die zentralen Elemente des Konzeptes AccSis. Die erzielbare Qualität des zu bestimmenden Risikos ist entscheidend abhängig von der Steuerbarkeit des Intensivierungsfaktors a , der zu der jeweiligen Zeitdrucklast konstant zu halten ist. Abweichungen durch nicht näher bestimmbare Unsicherheiten (Fehlervarianzen) führen vermutlich zu einer höheren Irrtumswahrscheinlichkeit bei der Verifikation des Risikos. Weiterer Zweifel an der praktischen Effektivität des Induktionsverfahrens kann aufgrund der Künstlichkeit der Arbeitsumgebung erhoben werden, da das Verfahren zur Zeitdruckinduktion keine de facto Bedrohung innerhalb einer Simulation bieten kann. Es ist ebenso sehr wohl vorstellbar, dass bestimmte Merkmale des Menschen zu einer Unempfänglichkeit des Zeitdrucks führen. Ebenso stehen möglicherweise die Verfahren des Flugbetriebs der effektiven Zeitdruckinduktion entgegen, da es i.d.R. einem zeitdruckmindernden Konzept folgt. Aus diesem Grund werden im Rahmen dieser Arbeit zwei Anforderungen an die Zeitdruckinduktionen gestellt, die für die resultierende Irrtumswahrscheinlichkeit nach dem letzten Verfahrensschritt von AccSis entscheidend sind.

- *Reproduktion des empfundenen Zeitdrucks*: Die Reproduktion steht in direkter Verbindung mit dem Merkmal der Steuerbarkeit. Sie ist entscheidend dafür, vordefinierte Zielzeitdrucklasten über die Zeit konstant halten zu können und somit gleiche Randbedingungen zu allen Stichprobenzeiten zur jeweilig gewählten Zeitdrucklast gewährleisten zu können. Eine gleichbleibende Intensivierung der Unfallwahrscheinlichkeit kann somit hervorgerufen werden.
- *Kalibrierung des Zeitdrucks*: Die Kalibrierung hat das Ziel, ein Maß der Zeitdrucklast gegenüber dem Designstress zu liefern. Der Designstress bemisst sich dabei an den Grenzwertgrößen t_{a0} und t_{n0} . Die Kenntnis über das Ausmaß des effektiv vermittelten Zeitdrucks würde es gestatten, den zeitdrucksensitiven Bereich systematisch zu identifizieren und abzutasten.

Die Schlüsselrolle zur Erfüllung dieser beiden Anforderungen spielt das Verfahren, welches Zeitdruck induzieren soll. Es ist derart auszulegen, dass der Mensch einen gleichmäßigen und kontrollierbaren Konkurrenzdruck und induzierten Zeitdruck erfährt. Die technische Umsetzung des Verfahrens und die dazugehörigen Experimentalbedingungen, die für eine effektive Zeitdruckinduktion bei gleichzeitiger Ausübung einer umfangreichen Primäraufgabe notwendig sind, sind bislang nicht systematisch untersucht worden. Die wichtigsten zu untersuchenden Gestaltungsfragen umfassen im Wesentlichen

- die effektive Vermittlung des Vorsprungs gegenüber dem Referenzoperator, dem „Rivalen“, die viel Spielraum zur freien Gestaltung lässt. Hier können z.B. visuelle und auch akustische Mittel eingesetzt werden.
- Der Einfluss und die Relevanz der Abwägung zwischen Risikobereitschaft und Schnelligkeit sind zu untersuchen.
- Das Bedrohungspotential der Bestrafung ist entscheidend für die Intensität, mit der der Zeitdruck empfunden wird. Die Umgebung eines Simulators macht es dabei schwierig, eine Bestrafungsmaßnahme einzurichten, die den Operator bei Überschreiten der zulässigen Maximalzeit realistisch bedrohen könnte.
- Die wichtigste Frage bleibt unzweifelhaft, wie hoch die Intensivierung der Unfallwahrscheinlichkeit maximal ausfallen kann. Sie ist ein Gradmesser für die Effektivität und die maximal zu erwartende Minderung der Irrtumswahrscheinlichkeit. Hier ist insbesondere das Zusammenwirken mit der primären Aufgabendefinition eine entscheidende Randbedingung, wenn diese z.B. durch eine konfliktfreie Verfahrensauslegung risikomindernd wirken soll und dem Induktionsverfahren somit zeitdruckmindernd entgegengerichtet ist.

Diese exemplarisch aufgeführten Fragestellungen machen deutlich, dass die Zusammenhänge zwischen der durch AccSis erzielbaren Qualität und die dafür notwendigen experimentellen Randbedingungen Gegenstand einer Untersuchung sein sollten. Das Erlangen von Grundlagenwissen hierüber kann zu Hinweisen führen, welche Bedingungen für die effektive Nutzung von AccSis für die Risikoanalyse vorliegen müssen und welche Möglichkeiten der Optimierung sich anbieten.

3 Experimentelle Studie

Das Kapitel 2 hat zwei Risikomodelle und das Zeitdruckinduktionsverfahren als Kernelemente des Konzeptes AccSis vorgestellt. Die Gültigkeit der unterstellten Zusammenhänge ist Voraussetzung für eine Minderung der Irrtumswahrscheinlichkeit und somit für das Erzielen einer Beschleunigung. Um die Eignung von AccSis für die retrospektive Risikoanalyse nachzuweisen, ist eine praktische Erprobung des Zeitdruckinduktionsverfahrens ein wichtiger erster Schritt und das Ziel dieses Kapitels. Die messbare Erhöhung der Unfallwahrscheinlichkeit wäre somit ein Nachweis für

- die Wirksamkeit des Verfahrens zur Induktion von Zeitdruck und
- die Gültigkeit des erweiterten Risikomodells.

Für den Fall, dass ein solcher Nachweis erbracht werden kann, stehen zwei Anforderungen an die damit verbundene Qualität im Fokus der Untersuchung. Diese Anforderungen beziehen sich auf die Kalibrierung und die Reproduktion, mit der das Zeitdruckinduktionsverfahren Zielzeitdrucklasten anfahren kann. Diese Anforderungen wurden im letzten Abschnitt des vorherigen Kapitels vorgestellt.

Die erwarteten Probleme bei der Übertragung von ALT sind sehr grundsätzlicher Natur. Der Grund hierfür ist beispielsweise, dass die zeitdruckinduzierte Unsicherheit in der Entscheidungsfindung und die gleichzeitige Kontrolle sich naturgemäß entgegenstehen. Allerdings kann der Widerspruch für den Fall gelöst werden, dass es sich hierbei um eine Kontrolle über das Ausmaß der Unsicherheit handelt. Als Analogie kann hier das Prinzip der Temperaturregelung dienen. Die thermische Energie kann als Grad der Unordnung des Stoffes erachtet werden, die dennoch hinsichtlich einer zu erreichenden Zieltemperatur kontrolliert werden kann. Die Schwierigkeit, die Unsicherheit der Arbeitsqualität von Menschen kontrollieren zu können, wurden in Abschnitt 1.5.1 bereits als bedeutendes Problem herausgestellt. Gegenüber technischen Produkten kann die menschliche Stressreaktion eine weitaus größere Varianz trotz Gleichhaltung aller umgebenden Arbeitsbedingungen aufweisen. Das Ausmaß dieser Varianz kann zu kleinen oder gar unmöglich zu messenden Effekten führen.

Die Nutzung des Zeitdrucks qualifiziert sich an dieser Stelle aufgrund der Eigenschaft, durch die dosierbare Vorgabe verfügbarer Zeiten gleichbleibende Stressbedingungen für den Menschen aufbauen zu können. Die induzierte Unsicherheit scheint beherrschbar und somit kontrollierbar. Wie soeben schon am Beispiel der Varianz angesprochen wurde, kann die Querwirkung von Störfaktoren eine effektive Induktion systematisch oder unkontrolliert verfälschen. Dies ist insbesondere kritisch, wenn Störfaktoren durch geeignete Versuchsbedingungen nicht randomisiert werden können oder systematisch zur gesetzten Last variieren. Wie durch Rastegary vorgestellt, kann z.B. die Wirkung des Zeitdrucks je nach Individuum (eiliger & nicht-eiliger Typus) unterschiedliche Leistungsfähigkeiten des Menschen abrufen (Rastegary & Landy, 1993). Aus diesem Grund ist es ebenso das Ziel dieses Kapitels, Faktoren und deren Bedingungen zu identifizieren, die den o.a. Anforderungen förderlich oder abträglich sind. Die Kenntnis über derartige Querwirkungen kann somit zur Evaluation des Verfahrens und somit der Steigerung der Effektivität dienen.

Für die Gewinnung erster belastbarer Aussagen zu den soeben aufgeführten Punkten soll die Erprobung im Rahmen einer experimentellen Pilotstudie durchgeführt werden. Diese Studie umfasst zu diesem Zweck eine prototypische Implementierung und einen praktischen Test des Verfahrens unter Laborbedingungen. Dabei beschränkt sich der Nachweis innerhalb des ersten Schritts auf Messbarkeit des Effektes von induziertem Zeitdruck bei Ausübung von Aufgaben einer spezifischen Arbeitsposition des Flugbetriebes. Die Ergebnisse sind aufgrund der Laborbedingungen und der damit verbundenen internen Validität eingeschränkt verallgemeinerbar. Die zu untersuchenden zeitdruckbedingten Effekte dieses Kapitels sind in der Übersicht vorgestellt:

1. *Sensitivität des Risikos zum induzierten Zeitdruck*: Die Abhängigkeit des Intensivierungsfaktors vom Zeitdruck ist durch die Variation der gesetzten verfügbaren Zeit zu ermitteln.

Neben diesem sollen Aussagen zu charakteristischen Kennwerten, Störvariablen und die dazugehörigen Randbedingungen gesammelt werden:

2. *Individualvarianz der Ergebnisse*: Der Einfluss von individuellen Personenmerkmalen, die einen Effekt auf die Leistungsfähigkeit und Kompensationsstrategien des Operateurs haben, können die jeweilige Effektivität des Verfahrens beeinflussen. Je nach Veranlagung des Operateurs könnte die gleiche Zeitdrucklast zu unterschiedlichen Stressreaktionen führen.

3. *Maximale Intensivierung der Unfallwahrscheinlichkeit*: Die größtmögliche Intensivierung ist ein charakteristischer Kennwert des Induktionsverfahrens. Die Kenntnis über diesen Wert könnte Auskunft geben, in welchem Maße die Rechtsensur gelöst werden kann.

4. *Wechselwirkung mit der Basislast*: Der induzierte Zeitdruck steht in einer Wechselwirkung mit der Basislast der Primäraufgabe. Die Kenntnis über das Ausmaß der Basislast als „Designstress“ ist Voraussetzung für das kalibrierte Ansteuern von Zielzeitdrucklasten und die Bestimmung des Risikos bei Designstress.

5. *Risikobereitschaft und Schnelligkeit*: Die Abwägung des Operateurs zwischen diesen Ansprüchen ist ein Personenmerkmal und stellt einen unbekanntem Störeinfluss auf den Intensivierungsfaktor dar.

Die Pilotstudie dient darüber hinaus dem Zweck,

6. Hypothesen für eine mögliche folgende Regelstudie entsprechend des Risikomodells zu entwickeln bzw. experimentell zu bestätigen und

7. die Datengrundlage für die Stichprobenplanungsgrößen einer möglichen folgenden Regelstudie zu gewinnen.

Um Aussagen zu den Punkten 1 bis 5 ermitteln zu können, wird das Kapitel wie folgt aufgebaut: Zunächst werden die relevanten Größen vorgestellt, unterschieden nach abhängigen und unabhängigen Variablen. Anschließend werden unter Zuhilfenahme des erweiterten Risikomodells des Konzeptes AccSis (Abschnitt 2.6.2.2) vorläufige Kausalhypothesen ergründet, die eine Stressreaktion der abhängigen Faktoren bei Variation der verfügbaren Zeit erwarten lassen. Als exemplarische Betriebsumgebung werden die Flugplatzkontrolle und die Arbeitsposition des Platzverkehrsloten am Flughafen Frankfurt am Main gewählt, die sicherheitsrelevante Aufgaben umfassen und dennoch klar abgrenzbare Entscheidungssituationen für die Beobachtung und Untersuchung bieten. Nach Definition der

Anforderungen an die Versuchsdurchführung wird das experimentelle Design vorgestellt, das die Aufgabendefinition, den Versuchsaufbau und die Verfahren zur Datengewinnung vorstellt. Die Anforderungen an das zu entwickelnde experimentelle Design orientieren sich an dem *Max-Kon-Min-Prinzip*, das grundlegende Grundsätze der Auslegung des Designs beschreibt (Kerlinger, 1964). Das Prinzip stellt Anforderungen an die Primärvarianz, Sekundärvarianz und Fehlervarianz, um mit Hilfe einer robusten Auslegung des Designs belastbare Aussagen aus den Messergebnissen ziehen zu können.

Die Implementierung des Zeitdruckinduktionsverfahrens *ComPerf* wird mit Hilfe eines softwarebasierten Fluglotsenagenten als Referenzoperator realisiert. Die innerhalb des Risikomodells definierte Produktivität wird definiert als die Leistung des Platzverkehrslotsen, den Flugverkehr effizient anzuweisen. Der Modellierung der operatorspezifischen Leistung kommt die wichtige Bedeutung zu, die Leistung des menschlichen Operators anhand von geeigneten Größen zu indizieren. Die Bestimmung der Leistungsmetrik dient schließlich dem Leistungsvergleich des menschlichen Operators mit dem Referenzoperator.

Für die Dimensionierung der Arbeitsumgebung werden zwei Klassen von Szenarien konzipiert: die Baseline und das Zeitdrucklastszenario. Hierzu passend werden Kalibrierungsverfahren entwickelt und angewandt, die einer geeigneten Bemessung der Verkehrslasten und einer Parametrierung geeigneter Zeitdrucklasten auf Basis des induzierten Zeitdrucks dienen. Anhand der Parameter können drei Szenariokonfigurationen begründet werden, nach denen die Versuchsplanung ausgelegt wird.

Die Daten werden nach Durchführung der Experimente hinsichtlich der Stressreaktionen auf ausgewählten Ebenen des Risikomodells analysiert. Die Ergebnisse dienen dazu, Aussagen zu den o.a. 7 Punkten treffen zu können. Im Vordergrund steht hierbei die Frage, ob eine Erhöhung der Unfallwahrscheinlichkeit anhand der gewählten Risikometrik feststellbar ist.

3.1 Das Konzept der Studie

3.1.1 Die Wahl der unabhängigen und abhängigen Größen

Der im vorhergehenden Abschnitt beschriebene Nachweis soll anhand der Beobachtung einer Reaktion bei gezielter Variation der verfügbaren Zeit t_a erbracht werden. Dies soll bei gleichzeitigem Ausschluss aller anderen Erklärungsmöglichkeiten gelingen. Dabei erfüllt die verfügbare Zeit die Funktion eines Regressors (unabhängige Größe) und die abhängige Größe die des Regressanden. Die Echtzeitsimulation bietet für den eindeutigen Nachweis eines Effektes den Vorteil, alle Randbedingungen außer der Zeitdrucklast kontrollieren und somit gleichhalten zu können. Die relevanten Größen der experimentellen Validierung werden wie folgt vorgestellt.

3.1.1.1 Die unabhängige Größe

- die verfügbare Zeit t_a ist eine dem Probanden maximal gesetzte Zeit zur Ausübung seiner Aufgaben gemäß dem definierten Induktionsverfahren in Abschnitt 2.5.4. Diese Größe wird aufgrund ihrer Funktion auch als *induzierte Zeitdrucklast* bezeichnet.

3.1.1.2 Die abhängigen Größen

- die Häufigkeit beobachteter sicherheitsrelevanter Ereignisse X_n (Risikometrik) pro Simulationslauf
- die Häufigkeit beobachteter Zeitfehler (ZF) pro Simulationslauf, die wie folgt definiert ist

$$\frac{t_n}{t_a} > 1 \quad (3.1)$$

- Leistungsmetrik (Produktivität) des menschlichen Operateurs
- der Leistungsvorsprung des menschlichen Operateurs gegenüber dem Referenzoperateur
- die Reaktionszeit des Probanden pro Entscheidung

3.1.1.3 Die zusätzlichen Beobachtungsgrößen

Die Befragung der Probanden während der Versuchsdurchführung soll zusätzliche Erkenntnisse über die Effektivität des Verfahrens liefern und dient der zusätzlichen Prüfung des vermittelten Zeitdrucks.

3.1.2 Die Kausalhypothesen

Die erwarteten Stressreaktionen sind bereits anhand des erweiterten Risikomodells hergeleitet worden. Bei zunehmender Verknappung der verfügbaren Zeit t_a werden die folgenden Stressreaktionen erwartet:

1. Zunahme der Anzahl sicherheitsrelevanter Ereignisse X_n pro Zeit,
2. Zunahme der Anzahl an Zeitfehlern ZF pro Zeit,
3. Zunahme der Leistung,
4. Abnahme der benötigten Zeit t_n und
5. Zunahme des Zeitbudgets TB

Die Stärke der Stressreaktion kann je nach gewählter Basislast unterschiedlich stark ausfallen. Diese Bereiche zeichnen sich mutmaßlich durch unterschiedliche Empfindlichkeiten aus oder durch keine Korrelation für den Fall der Sättigung bzw. der Überlast. Im zweiten Fall kann dies eine maximal mögliche Induktion unter extremen Wettbewerbsbedingungen sein, die keine Bedingungen für eine Steuerbarkeit bietet. Aus diesem Grund gelten die aufgeführten Kausalhypothesen für einen zeitdrucksensitiven Bereich. Dieser Bereich wurde in Abschnitt 2.6.2 vorgestellt und liegt im oberen Lastenbereich, oberhalb der Basislast (vgl. Abbildung 29 rechts).

3.1.3 Die Anforderungen an die Pilotstudie

3.1.3.1 Die Messung der Stressreaktionen

Die abhängigen Zielgrößen der Wirkungskette sind innerhalb der Echtzeitsimulation durch geeignete Messungen und Befragungen entsprechend Abschnitt 3.2.6 zu erfassen. Für die Bestimmung der unterstellten Effekte nach Induktion des Zeitdrucks sollen die Szenarien zwei Intensitäten der Zeitdrucklast auf Basis des Zeitdruckinduktionsverfahrens „Konkurrenzdruck“ umfassen. Dies stellt eine Primärvarianz her. Den Vergleich der

Ergebnisse beider Stufen liefert die Stressreaktion. Systematische Störeinflüsse sollen durch Anwenden des Ceteris-Paribus-Prinzips kontrolliert werden, die somit eine Minderung der Sekundärvarianz bewirken. Alle übrigen Fehlergrößen sollen mit Hilfe eines standardisierten Experimentaldesigns minimiert werden.

3.1.3.2 Die Aufgabendefinitionen

Innerhalb der Luftfahrt bietet sich eine Auswahl an Arbeitspositionen, die sich für den Nachweis potentiell eignen. Den Arbeitspositionen ist zu eigen, dass ihre Entscheidungen und Handlungen in hohem Maße sicherheitsrelevant sind. Eine Auswahl der Aufgabenbereiche ist beispielhaft in Tabelle 4 zusammengestellt.

Tabelle 4: Übersicht einer Auswahl an Arbeitspositionen mit direktem sicherheitsrelevantem Bezug zum Flugbetrieb

<i>Arbeitsposition</i>	<i>Aufgabendefinition</i>
Luftfahrzeugführer	ICAO Annex 2, Abschnitt 2.3.1 ff. (ICAO, 2005)
Streckenfluglotse/Anfluglotse	ICAO PANS-ATM Doc. 4444, Abschnitt 4.1 ff (ICAO, 2007b)
Platzverkehrsotse	ICAO PANS-ATM Doc. 4444, Abschnitt 7.1.1.1, mit den Verfahrensdefinition in Abschnitt 7.1.1.3 (ICAO, 2007b)
Rollverkehrsotse	ICAO PANS-ATM Doc. 4444, Abschnitt 7.1.1.1, mit den Verfahrensdefinition in Abschnitt 7.1.1.3 (ICAO, 2007b)
Vorfeldotse	ICAO Annex 14, Abschnitt 9.5 (ICAO, 2013a)

So ist insbesondere die verfahrenskonforme Durchführung das entscheidende Qualitätsmerkmal der Handlungen, für alle Beteiligten des Flugbetriebs einen erwartungskonformen und somit sicheren Ablauf zu gewährleisten.

3.1.3.3 Die Implementierung des Referenzoperators

Die Implementierung des Verfahrens zur Zeitdruckinduktion setzt voraus, innerhalb der Echtzeitsimulation über einen Referenzoperator zu verfügen, der die Aufgabenausführung unter einer vorgebbaren verfügbaren Zeit pro Entscheidung beherrscht. Dieses können angewandte Agentenmodelle sein, die befähigt sind, Entscheidungen auf taktischer als auch auf strategischer Ebene nach vordefinierten Verhaltensregeln zu treffen. Ebenso gehört hierzu die Aufgabe, zu priorisieren und zu sequenzieren, um Entscheidungssituationen im Betriebsablauf nach gegebenen Zielkriterien zu lösen. Der Vorteil gegenüber einem menschlichen Referenzoperator ist das gleichbleibende Verhalten des Agenten, das reproduzierbar ist und die Fehlervarianz somit minimieren soll.

3.1.3.4 Die Ausbildung der Probanden

Die Komplexität und Vielfalt der Aufgaben der in Abschnitt 3.1.3.2 vorgestellten Auswahl an Arbeitspositionen stellen enorme Forderungen an die Qualifikation des Operators. Um die Messung der Stressreaktion hinsichtlich mangelnder Fähigkeiten zur Ausübung der Fluglotsentätigkeiten zu schützen, soll eine entsprechende Ausbildung der Probanden die qualifizierten Aufgabenausführung innerhalb der Studie sicherstellen. Dies soll eine

Minderung der Sekundärvarianz bezwecken. Dies betrifft sowohl die Ebene des fachlichen Wissens als auch die handwerkliche Ebene. Hierzu sollen zwei Maßnahmen dienen:

- Die Qualität der Aufgabenausführung im simulierten Betrieb wird durch eine Vorauswahl nach Fähigkeit gesteigert. So soll das Verfahren zur Vorselektion insbesondere die Verfahrenstreue des Handelns gewichten, bei der nach einer zuverlässigen und gleichbleibenden Qualität der Aufgabenausführung zu selektieren ist. Diese Vorgehensweise entspricht der Ziehung einer geschichteten Zufallsstichprobe (eng. *Stratified Random Sampling*) zum Zweck der Varianzreduktion.
- Ein Training und Qualifikationstest wird vor dem Hauptversuch durchgeführt-

Aufgrund der Vorauswahl der Probanden ist diese Pilotstudie als *Quasi-Experiment* zu erachten, da eine Randomisierung der Personenmerkmale zum Zweck der Populationsvalidität nicht vorgenommen wird (Bortz & Döring, 2006). Die in dieser Studie unterstellte Validität bezieht sich allerdings auf eine Arbeitsposition, deren ausführender Operateur einer gesetzlichen Qualifikation zu entsprechen hat. In Deutschland ist dies durch die Verordnung *LuftPersV* geregelt, die auf Basis der *Verordnung (EU) Nr. 1178/2011* europäische Ausbildung und Lizenzierungsstandards für Luftfahrzeugführer definiert. Die aufgeführten Maßnahmen sind demnach ähnlich für die Auswahl qualifizierten Personals für die Ausführung von Flugkontrolldiensten. Trotz der Auswahl kann von einer Validität gegenüber der Gruppe der Fluglotsen ausgegangen werden.

3.1.4 Die Wahl der statistischen Tests

Für den Zusammenhangsnachweis kommen statistische Tests zur Anwendung, die einen positiven Befund anhand des P-Wertes¹⁵ belegen können. Die zur Analyse ausgewählten Tests werden im Folgenden vorgestellt.

1. *U-Ranktest nach Mann-Whitney*: Dieser Test prüft zwei nicht-parametrische, metrische und unabhängige Verteilungen anhand der Null-Verschiebung zweier Mediane \tilde{x} auf gleiche zentrale Tendenzen der Grundgesamtheit. Die hier vorliegende Untersuchung unterstellt zumeist eine Richtung der erwarteten Medianverschiebung als Alternativhypothese, für die ein einseitiger Test $H_0: \tilde{x}_1 \leq \tilde{x}_2$ oder $H_0: \tilde{x}_1 \geq \tilde{x}_2$ zur Anwendung kommt.
2. *T-Test*: Dieser Test prüft zwei standardverteilte metrische Verteilungen anhand der Mittelwertverschiebung auf Gleichheit.
3. *Einfache lineare Regression*: Dieser Test liefert das Bestimmtheitsmaß R^2 für die Beschreibung der Reaktion eines abhängigen Stichprobensatzes (Regressand) von einer gesetzten unabhängigen Variablen (Regressor) durch ein einfaches lineares Regressionsmodell, das den beobachteten Werten z.B. mit Hilfe der Methode der kleinsten Quadrate angepasst wird. Das lineare Regressionsmodell hat die Form

$$Y_i = a + b \cdot X_i + \epsilon, \quad (3.2)$$

mit der Verschiebung a (*Intercept*), der Steigung b (*slope*) und der Residue ϵ .

¹⁵ „Der P-Wert ist die Wahrscheinlichkeit, eine mindestens so große Prüfgröße wie die aus den Daten berechnete zu erhalten, wenn in Wirklichkeit die Nullhypothese gilt“ (Sachs, 1999, S. 188).

Das Bestimmtheitsmaß gibt dabei Auskunft über jenen Anteil der Varianz der beobachteten Daten, der durch das lineare Modell erklärt werden kann. Dies kann als Güte aufgefasst werden, die Abhängigkeit zwischen zwei Variablen durch eine lineare Gerade beschreiben zu können. Ein hohes Bestimmtheitsmaß gibt dabei Auskunft über kleine Abweichungen der abhängigen Variablen von der linearen Gerade.

4. *Korrelationsanalyse*: Dieser Test liefert den Korrelationskoeffizienten und das Bestimmtheitsmaß einer Korrelationsanalyse nach Pearson. Das Quadrieren des Korrelationskoeffizienten ergibt das Bestimmtheitsmaß R^2 .

Der kausale Zusammenhang gilt als nachgewiesen, wenn die alternative Hypothese nach Ablehnung der Null-Hypothese eine Unabhängigkeit des Regressanden ausschließt. Dies kann anhand des Steigungsfaktors b (Formel 3.2) wie folgt dargestellt werden:

$$H_1: b \neq 0 \quad (3.3)$$

Die zur Auswertung verwendeten Signifikanztests orientieren sich an den Signifikanzniveaus (Sachs, 1999) mit den P-Werten wie folgt:

1. einfach signifikant (*): $P \leq 5,0 \%$
2. sehr signifikant (**): $P \leq 1,0 \%$
3. hochsignifikant (***) : $P \leq 0,1 \%$

Die Signifikanztests werden mit Hilfe der Software *R*, Version 3.1.1 von *R Foundation* durchgeführt, deren Ergebnisse zur jeweiligen Prüfung dem Anhang beigelegt und innerhalb der Auswertung referenziert sind. Für die Nachvollziehbarkeit der Ergebnisse sind die Skripte als Dateien mit der Endung „.r“ im digitalen Anhang hinterlegt und können mit Hilfe des Befehls *source* innerhalb der Konsolenanwendung der Software *R* ausgeführt werden. Alle quantifizierten Werte werden mit einem Einheitenzeichen gemäß SI-Basiseinheitensystem angegeben.

3.2 Das experimentelle Design

3.2.1 Die Auswahl der Arbeitsposition

Aufgrund des sicherheitsrelevanten Zuständigkeitsbereiches wird die Pilotstudie am Beispiel der Arbeitsposition des Platzverkehrslotsen auf einem Großflughafen gewählt. Diese Position qualifiziert sich aufgrund einer umfangreichen Informationslage, die der Operateur für eine Entscheidungsfindung zu erwägen hat. Vorteilhaft sind die gut eingrenzbaeren Situationen der Entscheidungsfindung und Handlung für die Beobachtung und Messung der Zielgrößen. Ein besonderes Kriterium ist die Sicherheitsrelevanz der Handlungen in dieser Position. Im Zuständigkeitsbereich des Platzverkehrslotsen stehen gemäß (ICAO, 2016) die Unfallklassen der sogenannten *Runway Safety Accidents*. Zu diesen gehören im Wesentlichen:

- abnormaler Bodenkontakt (*Tail strike*)
- Vogelschlag
- Kollision auf der Start- und Landebahn (*Runway Collision*) aufgrund des Abkommens von der Start- oder Landebahn (*Runway Excursion*) oder aufgrund einer Störung auf der Start- oder Landebahn (*Runway Incursion*)

- Kontrollverlust auf dem Boden (*Loss of Control on Ground*)
- Kollision mit Hindernissen (*Collision with Obstacles*)

Der Anteil dieser Unfallklassen an der Gesamtanzahl von Unfällen im Zeitraum von 2006 bis 2010 lag in der kommerziellen Luftfahrt weltweit bei zwischen 50 % und 60 %. Der Zuständigkeitsbereich des Platzverkehrslotsen kann aus diesem Grund zurecht als Risikoschwerpunkt erachtet werden.

Die Sicherheitsrelevanz der Entscheidung von Platzverkehrslotsen wird deutlich, wenn die Unfälle hinsichtlich ihres Beitrages durch die Flugsicherung statistisch analysiert werden. Dies wurde beispielhaft durch eine Studie unter Verwendung des *Integrated Risk Pictures* (IRP) von Eurcontrol vorgenommen. Hierzu wurden Unfallberichte der europäischen kommerziellen Luftfahrt aus den Jahren 1998 bis 2004 und Vorfallberichte der Jahre 1990 bis 2002 hinsichtlich des Beitrages von Flugsicherungsaktivitäten zur Unfallentstehung ausgewertet (Eurocontrol, 2006b). Die Tabelle 5 zeigt die dazugehörige Risikoanalyse mit Ursachenbeitrag der Flugsicherung. Die Tabelle identifiziert das größte flugsicherungsbedingte Risiko bei der Kategorie *Runway Collision* mit einer relativen Häufigkeit von $6,3e-9$ pro Flug. Die Analyse zeigt weiterhin, dass der größten Anteil innerhalb dieser Kategorie durch startenden und landenden Verkehr im sogenannten „mixed mode“ verursacht wird (31 %). Die Ursachenanalyse ergab einen Anteil von 21 % durch inadäquate Anweisungen der Flugsicherung aufgrund einer verfehlten Wahrnehmung des drohenden Konflikts auf der Piste oder die Fehlbeurteilung der notwendigen Staffelung zwischen den Luftfahrzeugen.

Weiterhin konnten die weiteren Faktoren als unfallursächlich bestimmt werden (Danaher, 1980):

- unzureichende Abstimmung unter den Fluglotsen
- unpräzise Benutzung der Sprechfunk-Phraseologie
- vorübergehende Aufmerksamkeitsstörungen
- verwendung ungeeigneter Verfahren oder Technik
- fehlerhaftes Urteilsvermögen

Die statistischen Auswertungen zeigen gegenüber den anderen Phasen des Fluges ein sehr hohes Potential für Unfälle durch Fehlentscheidungen seitens des Platzverkehrslotsen. Das hohe Maß der Verantwortung über die Sicherheit durch die gewissenhafte Beobachtung des Betriebsumfeldes, korrekte Handlungen und verfahrenskonforme Entscheidungen sowie das hohe Eskalationspotential im Falle des Fehlers sind entscheidende Kriterien für die Wahl dieser Arbeitsposition für die Pilotstudie.

Neben dem Platzverkehrslotsen steht für die Pilotstudie ebenso die Luftfahrzeugführerposition zur Auswahl. Allerdings ist die damit verbundene Aufgabenvielfalt wesentlich höher und die mögliche Unfallentstehung ist durch eine hohe Anzahl an möglichen Ursachenfaktoren wesentlich vielschichtiger ausgeprägt. Die Modellierung einer geeigneten Leistungsmetrik der Produktivität wäre gegenüber dem Platzverkehrslotsen wesentlich umfangreicher. Ein Mehrwert der Wahl der Pilotenposition hinsichtlich der Untersuchungshypothesen dieser Pilotstudie ist demgegenüber aber nicht zu erwarten.

Für die Pilotstudie wird der Zuständigkeitsbereich des Platzverkehrslotsen auf die Positionen des Rollverkehrslotsen und des Vorfeldlotsen ausgedehnt. Dies dient dem Zweck, Handover-Verfahren zwischen den Zuständigkeitsbereichen im Untersuchungsraum und den damit verbundenen zusätzlichen Anforderungen an die Simulationsumgebung zu vermeiden. Eine derartige Erweiterung der Kompetenzen des Platzverkehrslotsen ist verfahrenstechnisch gedeckt, wenngleich dies auf kleineren Flughäfen üblich ist, wo kostensparende Argumente eine Rolle spielen.

Tabelle 5: Direkter Beitrag der Flugsicherung zu Unfällen von Luftfahrzeugen in der ECAC Region 2005 (Eurocontrol, 2006b)

<i>Unfallklasse</i>	<i>Unfallrate pro Flug</i>	<i>Relativer Anteil der Flugsicherheit</i>	<i>Unfallrate durch Flugsicherungsbeitrag pro Flug</i>
Zusammenstoß in der Luft (<i>Mid-air Collision</i>)	5,4e-09	64,5 %	3,5e-09
Kollision auf der Start- und Landebahn (<i>Runway Collision</i>)	3,3e-08	18,9 %	6,3e-09
Rollwegkollision (<i>Taxiway Collision</i>)	3,4e-09	9,2 %	3,1e-10
Bodenberührung mit einem nicht außer Kontrolle geratenen Luftfahrzeug (<i>Controlled Flight into Terrain</i>)	5,4e-08	1,5 %	8,3e-10
Wirbelschleppen-bedingter Unfall (<i>Wake Turbulence Accident</i>)	3,3e-09	6,9 %	2,3e-10
Kontrollverlust im Flug (<i>Loss of Control in Flight</i>)	1,3e-07	-	-
Kontrollverlust beim Start (<i>Loss of Control in Takeoff</i>)	4,8e-08	-	-
Kontrollverlust bei der Landung (<i>Loss of Control in Landing</i>)	6,4e-08	-	-
Strukturversagen	1,6e-08	-	-
Feuer und Explosion	2,2e-08	-	-
Gesamt	3,8e-07	2,9 %	1,1e-8

3.2.2 Die Auswahl der Risikometrik

Die Definition des Unfallereignisses als primäre Sicherheitsmetrik scheint zunächst plausibel im Sinne der Risikodefinition zu sein. Die innerhalb der Simulation notwendige Abbildung von Unfallereignissen erfordert allerdings die Berücksichtigung komplexer Vorgänge in der Simulation, die valides A-priori-Wissen verlangen. So ist umfassendes Wissen über geometrische Objektdimensionen und Wahrscheinlichkeiten für den Erfolg von Recovery-Maßnahmen, wie z.B. mögliche Ausweichmanöver, zu berücksichtigen. Dieses Problem wurde bereits in Abschnitt 1.4.3.1 in grundsätzlicherer Form vorgestellt.

Einen einfacheren Schluss auf die Unfallwahrscheinlichkeit liefert die Anzahl der eingetretenen Runway Incursion. Das Ereignis ist wie folgt definiert:

„Any occurrence at an aerodrome involving the incorrect presence of an aircraft vehicle or person on the protected area of a surface designated for the landing and take-off of aircraft.” (ICAO, 2007b, S. 32)

Wie schon in Abschnitt 2.1.1 vorgestellt, liegt die Begründung für diese Wahl in dem bereits erwiesenen Kausalzusammenhang zwischen dem Eintritt der Runway Incursion und einer korrelierenden Unfallwahrscheinlichkeit der Klasse Runway Collision (Eurocontrol, 2010b).

Die Häufigkeit der aufgetretenen Runway Incursion ist demnach eine direkte Folge der unsicheren Entscheidungsfindung sowohl seitens des Piloten als auch des Fluglotsen. Ist das simulierte Flugzeugführerverhalten aufgrund eines Agentenmodells ideal hinsichtlich der Verfahrenskonformität zu allen Situationen, kann ein Risiko seitens des Piloten ausgeschlossen werden. Das Risiko der Runway Incursion kann somit eindeutig auf die unsichere Entscheidungsfindung des Platzverkehrslotsen zurückgeführt werden. Durch die Division der Häufigkeit durch die Anzahl an Startfreigaben gemäß Abschnitt 2.1.3 entsteht die innerhalb dieser Pilotstudie verwendete Risikometrik *Runway Incursion-Rate*, kurz *RI-Rate*.

3.2.3 Die Definition der Primäraufgabe

Die Flugplatzkontrolle ist eine Dienstleistung, die anhand von Kontrollanweisungen und Freigabeerteilungen gemäß ICAO PANS-ATM Doc. 4444 (ICAO, 2007b),

1. Kollisionen zwischen Luftfahrzeugen und Hindernissen verhüten und
2. den geordneten und zügigen Ablauf des Flugverkehrs gewährleisten soll.

Der Zuständigkeitsbereich des Platzverkehrslotsen umfasst für die experimentelle Studie

- die Piste(n),
- den Endanflug nach dem *Final Approach Fix*¹⁶ in ca. 10 NM Entfernung vor der Pistenschwelle,
- den Abflug nach Abheben bis zu einem Wegpunkt in ca. 5000 ft Höhe in Richtung der Piste,
- die Auf- und Abrollwege der Piste und
- die Rollwege des Flughafens.

Aufgrund der Vorfeldkontrollfunktion des Platzverkehrslotsen wird der Zuständigkeitsbereich ergänzt durch

- das Vorfeld mit den zugehörigen Parkpositionen.

Die für diese experimentelle Studie verwendete Aufgabenvielfalt des Platzverkehrslotsen wird auf eine benötigte Kernkompetenz reduziert. Es handelt sich hierbei um eine Rationalisierung,

¹⁶ Der *Final Approach Fix* ist ein Navigationspunkt, der für Instrumentenverkehr den Anfangspunkt des Endanflug-Segments bildet.

die für die Untersuchung der Hypothesen legitim und ausreichend erscheint. Die Gründe hierfür sind

- eine reduzierte Menge an Anforderung für den notwendigen Funktionsumfang des Versuchsaufbaus
- ein reduzierter Trainingsaufwand zur Qualifikation von Novizenprobanden.

Tabelle 6: Auszug der zu erteilenden Freigaben und Anweisungen des Platzverkehrsleuten, Rollverkehrsleuten und Vorfeldleuten mit der maximal möglichen implementierten Schadensklasse

#	Freigabe/ Anweisung	Beschreibung	Schweregrad
1.	Startfreigabe	<ul style="list-style-type: none"> – Die Startfreigabe kann einem Luftfahrzeug erteilt werden, sobald Gewissheit über die ordentliche Staffelnung bei Beginn des Starts besteht (§7.9.3.1, ICAO PANS-ATM). – Für die Startfreigabe ist das Verfahren der reduzierten Pistenstaffelnung anzuwenden (§7.11.7, ICAO PANS-ATM). Für die Simulation innerhalb der Pilotstudie gelten die Vorschriften für Kategorie 3-Luftfahrzeuge von 2400m für <ul style="list-style-type: none"> – landende hinter startenden Luftfahrzeugen, – startende hinter startenden Luftfahrzeugen und – landende hinter landenden Luftfahrzeugen. – Ebenso ist die Wirbelschleppenstaffelnung für den Betrieb auf abhängigen Parallelpisten (§8.7.3.4, ICAO PANS-ATM) anzuwenden. 	Runway Collision
2.	Aufrollfreigabe	Luftfahrzeugen soll das Aufrollen und das Halten auf der Piste nicht erlaubt werden während ein anderes Luftfahrzeug eine Landung durchführt, bevor das landende Luftfahrzeug den Aufrollpunkt passiert hat (§7.6.3.1.3.2, ICAO PANS-ATM).	Runway Collision
3.	Kreuzungsfreigabe	Falls die Rollweganweisung über eine Piste führt, ist das Kreuzen anhand einer expliziten Freigabe anzuweisen (§7.6.3.1.1.2, ICAO PANS-ATM).	Runway Collision
4.	Rollfreigabe/ Rollweganweisung	Die Rollfreigabe sollte Anweisungen und geeigneten Informationen enthalten, um der Besatzung das Folgen der korrekten Rollwege bei gleichzeitiger Kollisionsvermeidung zu ermöglichen (§7.6.3.1.1.1, ICAO PANS-ATM).	
5.	Push Back	Die Freigabe liegt gemäß dem ICAO Annex 14, 9.5 Apron Management Service in dem Zuständigkeitsbereich des Flughafensbetreibers. In der Regel wird diese Freigabe nach Anfrage durch das Luftfahrzeug durch die Vorfeldkontrolle erteilt.	

Die für die Pilotstudie ausgewählten Aufgaben des Platzverkehrsleuten sind in Tabelle 6 zusammengefasst. Die Auflistung enthält ebenfalls die Pushback-Freigabe der Vorfeldkontrolle. Die Freigaben, die innerhalb dieser Studie ein Schadenspotential haben

sollen, sind in der Spalte *Schweregrad* mit der entsprechenden Klasse vermerkt. Das Auftreten jedweder Handlungssituation, in der der Operateur das Erteilen einer derartigen Freigabe erwägt, ist eine mögliche Gelegenheit für einen Unfall.

Alle übrigen Aufgaben und die dazugehörigen Handlungen sind nicht unmittelbar mit der Gelegenheit eines Unfalls verbunden. Sie dienen daher primär dem Zweck, einen Beitrag zur Herstellung einer geeigneten Basislast zu leisten. Ebenso soll die Qualität der Kontrolldienste sowie die Schnelligkeit der Handlungen zu einer messbaren Leistungsfähigkeit verbunden und dem Leistungsvergleich des Zeitdruckinduktionsverfahrens zugeführt werden. Unter die Rationalisierung der umfangreichen Menge an Aufgaben des Platzverkehrslotsen fallen ebenfalls zwei sehr bekannte und sicherheitsrelevante Freigaben bzw. Anweisungen (Tabelle 7), die im Folgenden diskutiert und deren Ausschluss vom experimentellen Design begründet werden.

- *Landefreigabe*: Obwohl die Entscheidung über das Erteilen der Landefreigabe eine hohe Kritikalität aufweist, ist sie nicht Bestandteil des Aufgabenumfangs. Begründet ist dies mit der Annahme, dass
 - i. durch den Anfluglotsen eine verfahrenskonforme Staffelung hergestellt wurde und
 - ii. der landende Verkehr generell Priorität vor kreuzendem, aufrollendem und startendem Verkehr bei Benutzung derselben oder einer kreuzenden Piste hat (§7.7, ICAO PANS-ATM).
- *Anweisen eines Durchstartmanövers/ Zurücknehmen von Freigaben*: Die Wahl dieser Handlungsoptionen durch den Platzverkehrslotsen stellt i.d.R. die zuletzt verbliebene Maßnahme zur Wiederherstellung der Verfahrenskonformität oder gar zur Vermeidung des Unfalls auf der Piste dar. Da im Bedarfsfall für eine solche Maßnahme der Eintritt einer Runway Incursion drohte oder bereits erfolgte, konnte das zu messende sicherheitsrelevante Ereignis bereits beobachtet werden. Die Beobachtung weiterer daraus resultierender Folgeereignisse wird für die Untersuchungshypothesen als nicht relevant erachtet.
- *Anlassfreigabe*: Das Verfahren zu Anlassfreigabe dient in der Praxis dazu, unnötige Staubbildung oder das Warten mit laufenden Motoren bzw. Triebwerken zu vermeiden. Diese Funktion wird innerhalb der Simulationsumgebung durch die Push-Back-Freigabe erfüllt.

Die Auflistung der Aufgaben in Tabelle 6 umfasst von den zahlreichen Verpflichtungen der Flugplatzkontrolle einen verhältnismäßig kleinen Ausschnitt. Aufgrund der Rationalisierung der Aufgabenmenge ist der entstehende simulierte Betrieb als eine Abstraktion des realen Vorbilds zu erachten, der dennoch die wesentlichsten sicherheitsrelevanten Wirkmechanismen zur möglichen Eskalation der Betriebssituation durch das unsichere Handeln des Operateurs bereitstellt. Der vorgestellte Aufgabenumfang erscheint somit für die Erprobung der Zeitdruckinduktion ausreichend geeignet.

Tabelle 7: Nicht berücksichtigte Freigaben und Anweisungen

#	Freigabe/ Anweisung	Beschreibung	Implementierte Schadensklasse
6.	Landefreigabe	Die Landefreigabe kann einem Luftfahrzeug erteilt werden sobald Gewissheit über die ordentliche Staffelnung beim Überfliegen der Pistenchwelle besteht (§7.10.2, ICAO PANS-ATM).	Runway Collision
7.	Durchstartanweisung/ Zurücknahme von Freigaben	Im Falle der Feststellung einer Runway Incursion [...] sollen geeignete Gegenmaßnahmen ergriffen werden (§7.4.1.4.1, ICAO PANS-ATM).	
8.	Anlassfreigabe	Verfahren für die Anlassfreigaben sollten eingerichtete werden um Staubildung auf den Rollwegen zu vermeiden. (§7.3.1.1.2, ICAO PANS-ATM).	

3.2.4 Der Surface Movement Manager

Der *Surface Movement Manager* ist das primäre Arbeitsinstrument des Versuchsaufbaus und eine Mensch-Maschine-Benutzerschnittstelle im Verbund mit einer umfassenden Simulationsinfrastruktur. Die Benutzerschnittstelle ist an eine Echtzeitsimulation angebunden, die Rollverkehr mit anfliegendem sowie abfliegendem Verkehr auf einem Flugplatz abbildet und die Gesten eines Bedieners dem simulierten Betriebsgeschehen zuführen kann. Die nun folgende Beschreibung leitet die wesentlichen Kernelemente dieser Infrastruktur ein.

3.2.4.1 Rollverkehrssimulator

Die Softwarearchitektur und die Regeln zur Generierung sowie Steuerung der Luftfahrzeugagenten ist in Anhang A.1 beschrieben. Die Verkehrslast wird mit Hilfe einer Menge an Luftfahrzeugen realisiert, die zeitgesteuert generiert werden kann. Der Eintrittszeitpunkt der Luftfahrzeuge in die Simulation wird anhand einer Gleichverteilung innerhalb eines definierbaren Zeitintervalls bestimmt, der durch eine zufallsbasierte Varianz überlagert wird. Im Intervall T können n Flugbewegungen wie folgt definiert werden:

$$\text{Verkehrsdichte} = \frac{n}{T} \quad (3.4)$$

3.2.4.2 Benutzerschnittstelle

Die Benutzerschnittstelle des Surface Movement Managers besteht aus zwei Visualisierungen (vgl. Abbildung 30, links):

- Bodengagedarstellung (Mode S)
- Luftlagedarstellung (Mode S und Multilateration)

Das System ist konform zu dem A-SMGCS Konzept (ICAO, 2004) mit den folgenden Freigaben des Platzverkehrsloten und Vorfeldloten, die in Abbildung 31 beispielhaft illustriert werden:

- Rollfreigabe (Abbildung 31 A und B)
- Aufrollfreigabe
- Startfreigabe (Abbildung 31 D)

- Freigabe zur Kreuzung der Piste (Abbildung 31 E)
- Freigabe zum Pushback (Abbildung 31 A)

Zusätzlich können Rollweganweisungen erteilt werden, die dem LFZ-Agenten mit Hilfe von Stiftgesten mitgeteilt werden können (Abbildung 31 C). Hierzu kommt der Eingabestift zum Einsatz (Abbildung 30, rechts).

Die Information über die Gesamtroute und auch die Wegstrecke bis zum nächsten freigabepflichtigen Punkt wird dem Operateur nach einer Stiftgeste angezeigt. Die Funktionalität entspricht somit dem Level 3 nach Eurocontrol-Implementierungslevel (Eurocontrol, 2010a):

- Überwachung (Surveillance)
- automatisierte Detektion von Konflikten (Control)
- Anweisung von Routen (Route Planning)
- Freigabeerteilung an freigabepflichtigen Punkten (Guidance)

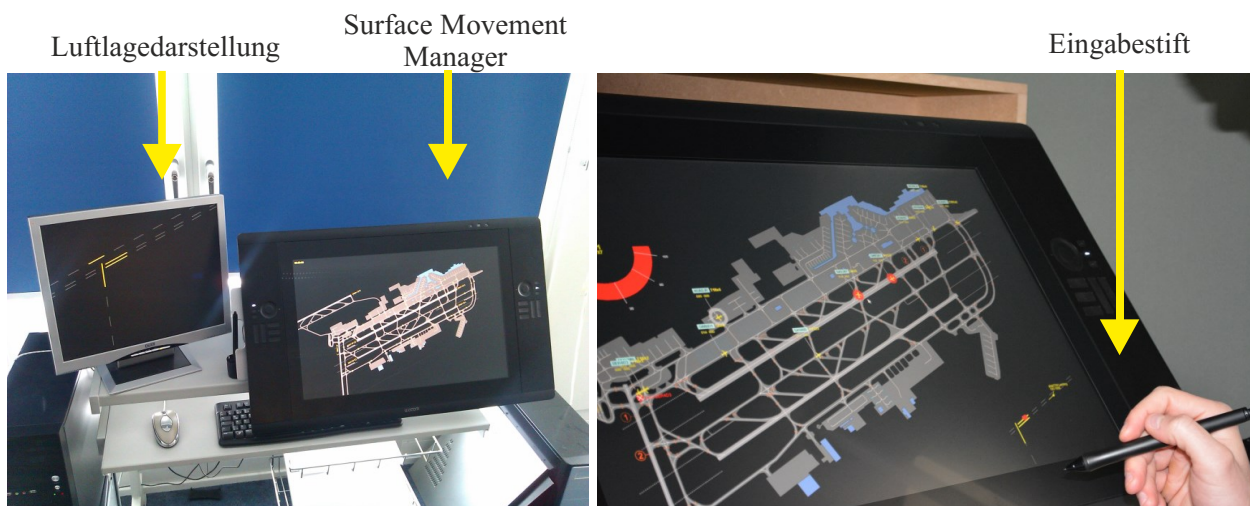


Abbildung 30: Der Surface Movement Manager

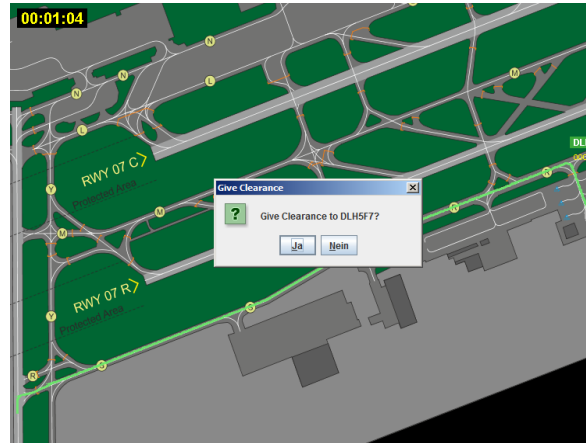
3.2.5 Die Implementierung des Verfahrens „Konkurrenzdruck“

3.2.5.1 Das Prinzip

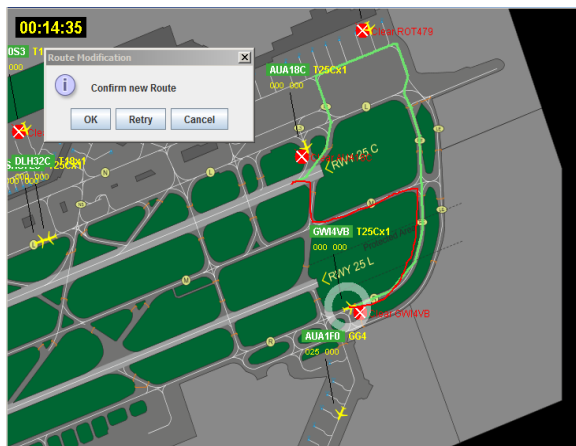
Das Verfahren zur Induktion von Zeitdruck wird für die Pilotstudie entworfen und implementiert. Zentrales Element ist die Messung der Referenzleistung eines softwarebasierten Fluglotsenagenten als Referenzoperator. Der Vorteil der Nutzung eines derartigen Agenten ist die Fähigkeit der Agentensimulation als Schnellzeitsimulation berechnet werden und reproduzierbares Verhalten erwarten zu können. Die Berechnungszeit für die Simulation eines vollständigen Szenarios ist somit abhängig von der maximalen Rechenleistung der technischen Simulationsinfrastruktur und der Rechenintensität der Szenarien. Letzteres ist vornehmlich abhängig von der Anzahl an Luftfahrzeugen und der Länge der Rollwege sowie Ab- und Abflugstrecken innerhalb der simulierten Betriebsumgebung.



A: Rollweganweisung Romeo in Richtung 18



B: Rollwegfreigabe



C: Kreuzen des Verfahrensschutzraums von 25L



D: Startfreigabe auf der 25C



E: Runway Incursion zwischen kreuzenden und startenden LFZ auf 25C



F: Runway Incursion zwischen zwei startenden LFZ auf 18 und 25C

Abbildung 31: Auswahl an möglichen Situationen im simulierten Flugbetrieb am Surface Movement Manager

Die Herstellung gleicher Anfangs- und Randbedingungen für den Fluglotsenagenten schafft einen Referenzbetrieb, der geeignete Voraussetzungen für den fairen Leistungsvergleich im Wettbewerb der Operateure bietet. Das Betriebsgeschehen kann aufgrund der unterschiedlichen gewählten Reihenfolgen und Zeitpunkte der Freigabeerteilungen z.B. zur Bildung der Abflugsequenz eine operateurspezifische Entwicklung nehmen. Je nach Effizienz der Freigaben und Anweisungen können die Prozesszeiten der zu kontrollierenden

Luftfahrzeuge variieren. Eine Bestimmung der Effizienz des menschlichen Operateurs im Wettbewerb mit dem Referenzoperator gelingt dabei durch einen Vergleich beider „Welten“ anhand einer geeigneten Leistungsmetrik.

Zur Indikation der Leistung bieten sich eine Vielzahl von Größen im fortlaufenden Betrieb an. Die Kriterien dafür, was als „geeignet“ erachtet werden kann, ist in diesem Abschnitt Gegenstand einer intensiven Diskussion. Die Anforderung ist es, dass der Proband die Regeln des Wettbewerbs plausibel und einfach nachvollziehen können soll. Hierzu gehört eine unmissverständliche Erklärung über den Begriff der Effizienz und welches Verhalten nach Leistungsvergleich belohnt oder bestraft wird. Ein Leistungsmodell des Platzverkehrslotsen soll den notwendigen Überblick über mögliche Kandidaten herstellen und somit eine Argumentationsgrundlage für die Wahl einer geeigneten Leistungsmetrik liefern.

Die Herstellung gleicher Anfangs- und Randbedingungen für den Fluglotsenagenten schafft einen Referenzbetrieb, der geeignete Voraussetzungen für den fairen Leistungsvergleich im Wettbewerb der Operateure bietet. Das Betriebsgeschehen kann aufgrund der unterschiedlichen gewählten Reihenfolgen und Zeitpunkte der Freigabeerteilungen z.B. zur Bildung der Abflugsequenz eine operateurspezifische Entwicklung nehmen. Je nach Effizienz der Freigaben und Anweisungen können die Prozesszeiten der zu kontrollierenden Luftfahrzeuge variieren. Eine Bestimmung der Effizienz des menschlichen Operateurs im Wettbewerb mit dem Referenzoperator gelingt dabei durch einen Vergleich beider „Welten“ anhand einer geeigneten Leistungsmetrik.

Zur Indikation der Leistung bieten sich eine Vielzahl von Größen im fortlaufenden Betrieb an. Die Kriterien dafür, was als „geeignet“ erachtet werden kann, ist in diesem Abschnitt Gegenstand einer intensiven Diskussion. Die Anforderung ist es, dass der Proband die Regeln des Wettbewerbs plausibel und einfach nachvollziehen können soll. Hierzu gehört eine unmissverständliche Erklärung über den Begriff der Effizienz und welches Verhalten nach Leistungsvergleich belohnt oder bestraft wird. Ein Leistungsmodell des Platzverkehrslotsen soll den notwendigen Überblick über mögliche Kandidaten herstellen und somit eine Argumentationsgrundlage für die Wahl einer geeigneten Leistungsmetrik liefern.

3.2.5.2 Die Modellierung der Leistung der Flugplatzkontrolle

Grundlegend entspricht der hier verwendete Begriff der Leistung einer Arbeitsleistung P , die von der Zeit t und der Qualität Q gemäß Formel 2.19 nach (Bubb, 2005) abhängig ist. In den Aufgaben der Flugplatzkontrolle gemäß ICAO PANS-ATM Doc. 4444, §7.1.1.1,

- den sicheren und
- geordneten sowie zügigen

Ablauf des Flugverkehrs im Zuständigkeitsbereich zu gewährleisten, lassen sich diese Ansprüche nach Leistung, Qualität und Schnelligkeit wiederfinden. Das erweiterte Risikomodell scheint aufgrund der Freiheit des Operateurs zur Abwägung zwischen Risikobereitschaft und Schnelligkeit für die Flugplatzkontrolle grundsätzlich anwendbar. Die notwendige Anpassung des erweiterten Risikomodells betrifft dabei im Wesentlichen die Produktivität, welche die Basis für den Vergleich der Operateure im Wettbewerb darstellt. Das zu entwickelnde Leistungsmodell soll den Zweck erfüllen, die menschliche Arbeitsleistung

des Platzverkehrslotsen anhand von Faktoren zu beschreiben und die Leistung hinsichtlich der erbrachten Flugplatzkontrolle abzuschätzen.

Für die Modellierung der Arbeitsleistung des Platzverkehrslotsen werden innerhalb dieses Abschnittes geeignete Parameter definiert, die die beiden soeben genannten Ansprüche abbilden. Es handelt sich hierbei um eine starke Abstraktion der Arbeitsleistung durch Anwendung vereinfachter Faktoren, die eine oberflächliche „Modelltiefe“ bieten. Es ist dennoch davon auszugehen, dass das Maß der gewählten Dekomposition der Faktoren für die Beschreibung der Flugplatzkontrolle zweckdienlich und für die Argumentation einer geeigneten Leistungsmetrik ausreichend ist.

Risikobereitschaft

Die Risikobereitschaft wird für diese Studie definiert als ein Merkmal, eine einzelne Entscheidungssituation hinsichtlich eines individuellen akzeptierbaren Risikos zu lösen. Bei Auftritt einer sicherheitsrelevanten Entscheidungssituation beschreibt der *Qualitätsfaktor* $Q_{konfliktfrei}$ ein Maß der Sicherheit, welches die vorliegende Risikobereitschaft zur jeweiligen Entscheidung berücksichtigt. Die Indikation kann beispielsweise anhand einer Komplementärwahrscheinlichkeit der Unfallwahrscheinlichkeit erfolgen:

$$Q_{konfliktfrei} = 1 - p. \quad (3.5)$$

Ein Alternativansatz ist der Kehrwert mit

$$Q_{konfliktfrei} = 1/p. \quad (3.6)$$

Die darauf aufbauende *Sicherheitsleistung* $P_{konfliktfrei}$ der Flugplatzkontrolle setzt die erreichte Qualität in Relation zu einem Zeitraum mit mehreren Gelegenheiten. Dies folgt der Definition:

„the level of safety performance is normally represented by safety indicators related to high-consequence outcomes (such as accident and serious incident rates)“ (ICAO, 2013c, S. 78).

Als hierzu passende Risikometrik wurde in Abschnitt 3.2.2 die *RI-Rate* definiert, die einen Schluss auf die Unfallwahrscheinlichkeit zulässt.

Schnelligkeit

Der Anspruch nach Schnelligkeit (bzw. hier auch Effizienz) des Flugbetriebs ist eine Zusammensetzung aus

1. der Qualität der Freigabe- und Anweisungserteilung, die Verkehrsbewegung möglichst zügig durch den Zuständigkeitsbereiches zu führen, und
2. der damit verbundenen notwendigen Zeit, die durch den Operateur für die Optimierung aufgewendet wird.

Eine hohe Qualität zeichnet sich dabei durch die Fähigkeit des Platzverkehrslotsen aus, die einzelne Verkehrsbewegung in einem Gesamtbetrieb geschickt zu priorisieren und zu sequenzieren (*P&S*). Der Qualitätsfaktor $Q_{P\&S}$ ist aus diesem Grund definiert als ein Maß, die

Bewegungszeit zum Durchqueren des Zuständigkeitsbereiches einer Bewegung hinsichtlich eines möglichen Minimums optimiert zu haben. Der Faktor ist im hohen Maße abhängig von der operatorspezifischen Fähigkeit, mit den verfügbaren Kontrollmöglichkeiten unter den gegebenen Ressourcen des simulierten Flughafenbetriebes, wie z.B. Rollwege, inklusive den Rollgassen, Kreuzungsbereichen, den Pisten und der gegenwärtigen Verkehrslage, die einzelne Verkehrsbewegung zeitsparend und flüssig zu führen.

Die *Leistungsfähigkeit* $P_{P\&S}$ der Flugplatzkontrolle setzt die erreichte Qualität in Relation zur notwendigen Zeit des Operators, die für die Kontrolle der Einzelbewegung aufgebracht werden muss:

$$P_{P\&S} = \frac{Q_{P\&S}}{t_n}. \quad (3.7)$$

Eine Möglichkeit zur Messung der Leistung $P_{P\&S}$ ist die Zeit zum Durchqueren des Zuständigkeitsbereiches einer kontrollierten Verkehrsbewegung: die *Aktivzeit* t_{aktiv} (Abbildung 32, links).

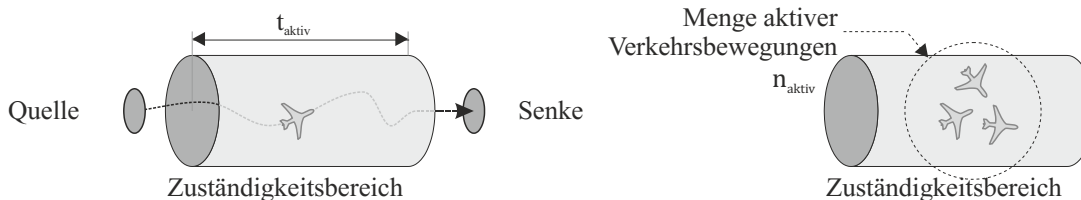


Abbildung 32: Die Aktivzeit als verbildlichte Reisezeit von der Quelle bis zur Senke (links). Menge der aktiven Verkehrsbewegungen (rechts)

Die folgenden Summanden definieren die Aktivzeit:

- Summe aller Haltezeiten
- Summe der Bewegungszeiten

Die Haltezeit setzt sich wiederum aus den folgenden Zeiten je freigabepflichtigem Punkt zusammen:

- Die *Reaktionszeit des Operators* t_r ist definiert durch die Zeit von der Freigabeanfrage durch die Verkehrsbewegung bis zum Zeitpunkt der Freigabeerteilung.
- Die *Verzögerungszeit* $t_{konfliktfrei}$ zu allen sicherheitsrelevanten freigabepflichtigen Punkten, definiert als eine notwendige Zeit der Verzögerung zur Herstellung bzw. Gewährleistung eines konfliktfreien Flugbetriebs.

Die Reaktionszeit wird als eine messbare und direkte Folgegröße der benötigten Zeit erachtet, die sich zur Indikation qualifiziert.

- Die Gesamthaltezeit resultiert für die Verkehrsbewegung somit zu

$$\sum_{\text{Freigaben}} (t_r + t_{konfliktfrei}). \quad (3.8)$$

- Die *Bewegungszeit* $t_{bewegung}$ ist die Summe der benötigten Zeiten von z.B. den folgenden Prozessen:

1. Pushback (für ausgehende Verkehrsbewegungen)
2. Rollen
3. Auf- oder Abrollen
4. Starten oder Landen
5. Anflug oder Abflug

Dem Platzverkehrslotsen bieten sich hier insbesondere bei den Punkten 2, 3 und 4 Spielräume, anhand von Sequenzieren und Priorisieren die Zeiten zu minimieren und die Qualität $Q_{P\&S}$ zu steigern. So kann z.B. die Reihenfolge der Starts und Landungen genutzt werden, um den Beschränkungen von Pistenabhängigkeiten optimal aufzulösen und eine maximale Kapazität abzugewinnen.

- Die Aktivzeit einer Verkehrsbewegung resultiert schließlich zu

$$t_{aktiv} = \sum_{Freigaben} t_n \left(t_x + \frac{t_{konfliktfrei}}{f(Q_{P\&S})} + t_{bewegung} \right) \quad (3.9)$$

Diese Zeit sinkt, wenn die Qualität $Q_{P\&S}$ steigt oder die notwendige Zeit für die Entscheidungsfindung sinkt. Dieser Zusammenhang entspricht der Definition der Leistung $P_{P\&S}$ Formel 3.7. Aus diesem Grund ist die Aktivzeit als geeigneter Indikator zu erachten, der mit der Leistung des Operateurs korreliert. Die soeben gezeigte Herleitung der Aktivzeit gilt für jede Verkehrsbewegung. Die Messung der Aktivzeit kann somit als Stichprobe im Rahmen einer Leistungsmessung dienen.

Die Bestimmung der Leistung einer Einzelbewegung berücksichtigt allerdings nicht die Fähigkeit der Platzverkehrslotsen, eine einzelne Verkehrsbewegung in einem Gesamtprozess zu priorisieren und zu sequenzieren. Wie zu Beginn dieses Abschnitts erläutert ist stattdessen eine Leistungsmetrik gefordert, die die Effizienz der einzelnen Verkehrsbewegung zu einer Gesamteffizienz aller Verkehrsbewegungen eines vordefinierten Zeitraums zusammenfasst.

Ausgangspunkt dieser Gesamteffizienz ist die Kontrolle des simulierten Flugverkehrs bei Generierung einer Verkehrsdichte mit einer Anzahl an Verkehrsbewegungen n pro Zeitintervall gemäß Formel 3.4,

$$\frac{n_k}{T} \quad (3.10)$$

mit den Intervallschritten der Reihe $k = 0,1,2,3, \dots, k \in \mathbb{N}$ und der Intervalldauer T . Dies ist konform zu der Flugplandefinition, beschrieben in Abschnitt 3.2.4.1, die den Anfang und das Ende des Intervalls mit einer Anzahl an Verkehrsbewegungen definiert.

Auf Basis der Aktivzeit versprechen zwei Vorschläge die Beschreibung eines betrieblichen Leistungsmaßes:

- *Arithmetischer Mittelwert der Aktivzeit*: Die Bildung des arithmetischen Mittelwertes der Aktivzeit über n kontrollierte Verkehrsbewegungen innerhalb eines vorgegebenen

Zeitraums liefert ein Maß der mittleren Leistungsfähigkeit des Platzverkehrslotsen, die betreffenden Verkehrsbewegungen effizient zu führen:

$$\frac{\sum_n t_{aktiv}}{n} = \hat{t}_{aktiv} \quad (3.11)$$

- *Anzahl aktiver Verkehrsbewegungen:* Diese Größe gibt die Anzahl derer wider, die gemäß der Verbildlichung durch Abbildung 32, rechts, zu einem bestimmten Zeitpunkt t im Zuständigkeitsbereich des Platzverkehrslotsen aktiv sind. Sie ist definiert als eine Differenz zwischen den generierten $n_{generiert}$ und abgeschlossenen $n_{abgeschlossen}$ Verkehrsbewegungen zu

$$n_{aktiv} = n_{generiert} - n_{abgeschlossen} \cdot \quad (3.12)$$

Die Anzahl $n_{generiert}$ ist die Summe der bis zum Zeitpunkt t generierten Verkehrsbewegungen gemäß Formel 3.10. Für den Intervallschritt $m \in k$ zum Zeitpunkt $(m + 1) \cdot T \cdot > t \geq m \cdot T$ ergibt sich die Anzahl zu

$$n_{generiert} = \sum_{k=0}^{m \leq t/T} n_k \cdot \quad (3.13)$$

Die Anzahl der abgeschlossenen Verkehrsbewegungen $n_{abgeschlossen}$ hat einen Leistungsbezug aufgrund des Zusammenhangs mit der Aktivzeit. Unter Annahme einer stochastischen Verteilung der Aktivzeiten mit noch unbekannter Verteilungscharakteristik kann der Zusammenhang zwischen der Verteilung von t_{aktiv} und n_{aktiv} grafisch nachvollzogen werden. Ein mögliches Verteilungsmodell der Aktivzeit ist die Weibullverteilung, die für die Beschreibung der Verteilung von Rollzeiten am Flughafen Frankfurt a. Main verwendet werden kann (Günther, Hildebrandt, Fricke, & Strasser, 2011). Abbildung 33 illustriert dies anhand eines Verteilungsbeispiels einer Wahrscheinlichkeitsdichte für die Aktivzeit einer Verkehrsbewegung (Abbildung 33, links). Eine minimale Aktivzeit $t_{aktiv\ min}$ existiert aufgrund von notwendigen Mindestrollzeiten und Zeiten des Startvorgangs und Endanflugs.

Die Anzahl aktiver Verkehrsbewegungen kann dabei unter Verwendung der kumulativen Verteilung der Aktivzeit veranschaulicht werden (Abbildung 33, rechts). Die Darstellung zeigt die Intervallzeiten auf der Zeitachse und die Normierung der Verteilung auf n_k . Der Kurvenverlauf verdeutlicht anhand des dunkelgrauen Bereiches die zunehmende Anzahl an erwarteten abgeschlossenen Verkehrsbewegungen bei zunehmender Aktivzeit. Die Anzahl der noch nicht abgeschlossenen Verkehrsbewegungen im hellgrauen Bereich stellt die Menge aktiver Verkehrsbewegungen dar. Eine abnehmende Anzahl aktiver Verkehrsbewegungen lässt bei konstanter Generation n_k den Rückschluss auf eine Verkleinerung der Aktivzeit zu. Demnach indiziert eine kleinere Anzahl eine höhere Leistungsfähigkeit.

Die innerhalb dieses Abschnitts hergeleiteten Zusammenhänge sind in Abbildung 34 als Wirkungsverkettung dargestellt und zusammengefasst. Das Modell stellt die Anzahl aktiver Verkehrsbewegungen als Maß der Produktivität der Flugplatzkontrolle dar. Die zwei Wirkungszweige der Qualität und der Zeit sollen die Kompromissfindung des Menschen verdeutlichen, die Arbeitsleistung je nach situativen Erfordernissen zwischen Schnelligkeit und Genauigkeit aufteilen zu können. Die Summenbildung der Aktivzeit über alle elementaren Prozesse verdeutlicht die Freiheitsgrade, gleiche Aktivzeiten durch Schnelligkeit als auch durch geschickte Optimierung erhalten zu können.

3.2.5.3 Die Arbeitsbedingungen der Operateure

Die Funktionen des Agenten sind gemäß der primären Aufgabendefinition des Abschnitts 3.2.3 implementiert. Dabei ist der Agent befähigt, den Verfahren zu jeder Freigabeerteilung vollständig zu entsprechen, um konfliktfreie Entscheidungen zu treffen. Dieses gelingt durch einen Kollisionsverhütungsalgorithmus, der jede Gelegenheit zur Freigabeerteilung mit Hilfe von Prognosedaten hinsichtlich der Verletzung der Pistenstaffelung evaluiert. Es entsteht eine verfahrenskonforme Nutzung der Piste, die als vollständig risikofrei angenommen werden kann. Die Arbeitspriorisierung, die die simultane Kontrolle mehrerer Bewegungen umfasst, erfolgt parallel und unabhängig je Bewegung. Dabei werden die durch die Verkehrsbewegungen gestellten Freigabebeanfragen gleichgewichtet priorisiert und durch den Agenten abgearbeitet.

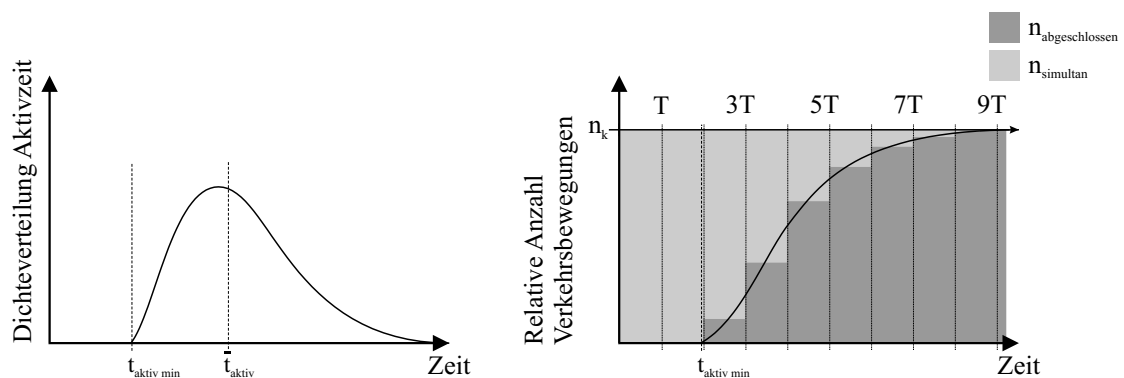


Abbildung 33: Wahrscheinlichkeitsdichte der Aktivzeitverteilung (links). Kumulative Verteilung der Aktivzeit (rechts)

Die Rollweganweisungen des Agenten entsprechen der gesetzten Standardroute, die durch den Surface Movement Manager nach einem Wegekostenoptimum ermittelt wurde. Das Verfahren zur Ermittlung der Standardroute wurde bereits in Abschnitt 3.2.4.1 vorgestellt. Das dynamische Modifizieren der Rollweganweisung während des Rollvorgangs ist konzeptionell nicht implementiert. Dies bedeutet, dass der Agent z.B. auf aktuelle Verkehrsströme und Staubildungen keine optimierten Rollweg ad hoc ermitteln und anweisen kann.

Die Arbeitsgeschwindigkeit des Agenten ist parametrisiert durch eine Reaktionszeit von der Freigabebeanforderung bis zum frühesten Zeitpunkt der Freigabeerteilung. Da es sich hier um die zu setzende Zeitdrucklast des Zeitdruckinduktionsverfahrens handelt, sollte sie sich terminologisch von der menschlichen Reaktionszeit abgrenzen. Hierzu wird das Synonym

der *Latenzzeit* mit dem Formelzeichen t_L verwendet. Die tatsächliche Freigabeerteilung $t_{freigabe}$ kann aufgrund der Intervention des Kollisionsverhütungsalgorithmus variieren. Der Algorithmus gewährleistet die vorgeschriebene Pistenstaffelung durch das Verzögern um eine Zeit $t_{konfliktfrei}$. Die Zeit von der Freigabeanforderung eines Luftfahrzeugs bis zur tatsächlichen Freigabeerteilung kann somit größer als t_L , zu

$$t_{freigabe} = t_L + t_{konfliktfrei} , \quad (3.14)$$

ausfallen. Somit ist die parametrierbare Latenzzeit t_L ein bestimmender Faktor für die Leistung des Fluglotsenagenten. Die Reihenfolge der Freigaben orientiert sich an der Reihenfolge der eingegangenen Anfragen zu allen freigabepflichtigen Punkten (*first-come-first-serve*).

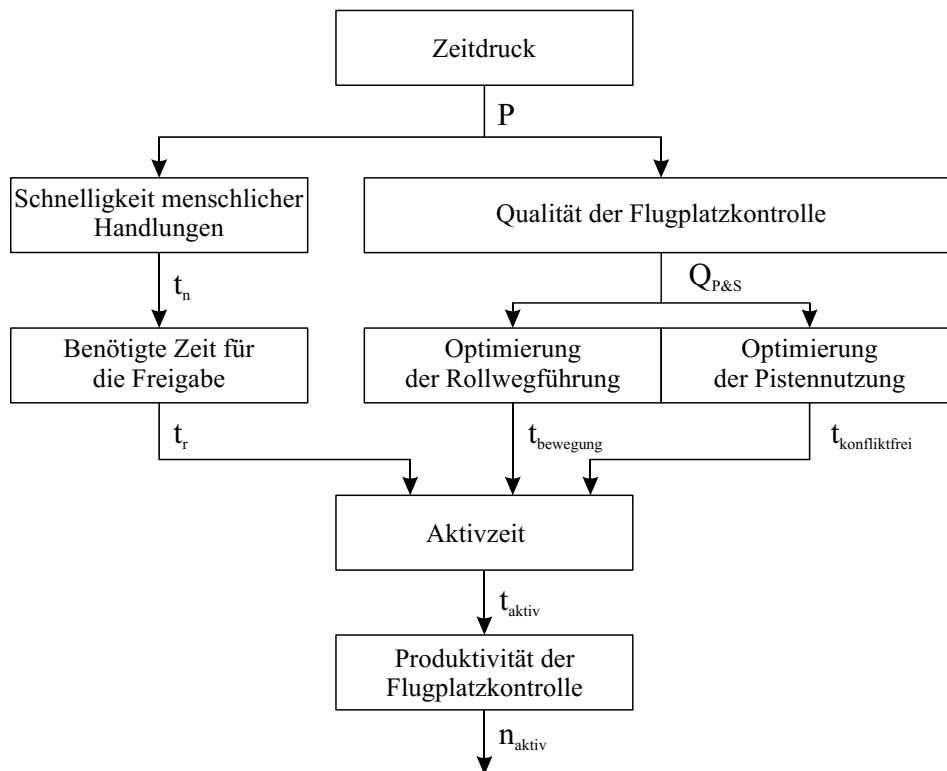


Abbildung 34: Das Leistungsmodell verbindet die Zeit der Einzelentscheidung und Qualität zu einer Anzahl aktiver Verkehrsbewegungen

Der menschliche Operateur hingegen entscheidet bei simultaner Kontrolle mehrerer Verkehrsbewegungen auf einer aggregierten Prozessebene, die das kontinuierliche Erfassen, Bewerten und Sortieren verteilter Verkehrsbewegungen im Zuständigkeitsbereich umfasst. Die Gegenüberstellung der Operateure verdeutlicht die konzeptionellen Gegensätze ihrer Arbeitsweise (Tabelle 8). Durch das streng vorbestimmbare Verhalten des Agenten kann ein Maximum an Sicherheit erreicht werden, das aber ansonsten wenig Potential zur dynamischen Optimierung des Gesamtablaufes erwarten lässt. An dieser Stelle könnte der menschliche Operateur einen Wettbewerbsvorteil erlangen, der insbesondere die Größe $Q_{P\&S}$ betrifft.

Tabelle 8: Qualitäten und Zeit der konkurrierenden Operateure als Faktoren der Ausgangsleistung in der Gegenüberstellung

<i>Operateur</i>	$Q_{P\&S}$	$Q_{\text{konfliktfrei}}$	t_{freigabe}
Fluglotsenagent	First Come - First Serve (keine Optimierung der Pistenbenutzung) keine Rollwegoptimierung	vollständig konfliktfrei	$t_L + t_{\text{konfliktfrei}}$
menschlicher Operateur	Planung der Pistenutzung Rollwegoptimierung, indiziert anhand der Anzahl von Rollwegmodifikationen n_{rm}	Risikobehaftet aufgrund der Unsicherheit menschlicher Handlungen	$t_r + t_{\text{konfliktfrei}}$

3.2.5.4 Die Bestimmung der Leistungsmetrik

Die konzeptionellen Unterschiede in den Arbeitsbedingungen führen zu der Notwendigkeit, eine geeignete Leistungsmetrik zur Indikation von $P_{P\&S}$ zu ergründen, die die Leistungsfähigkeit beider Operateure widerspiegeln kann. Das Leistungsmodell des Platzverkehrsloten liefert hierzu zwei interessante Kandidaten, die in Tabelle 9 anhand von Argumenten gegenübergestellt sind.

Tabelle 9: Gegenüberstellung von zwei vorgeschlagenen Leistungsmetriken für den Leistungsvergleich

<i>Merkmal</i>	<i>arithmetisches Mittel der Aktivzeit</i> (\hat{t}_{aktiv})	<i>Anzahl aktiver Verkehrsbewegungen</i> n_{aktiv}
Echtzeitbestimmung	Ja	Ja
einbezogene Verkehrsbewegungen	alle Verkehrsbewegungen, die abgeschlossen sind.	alle aktiven Verkehrsbewegungen
Vergleichbarkeit	Mittelwertvergleich gemäß Formel 3.11	Differenzbildung der Anzahl gemäß Formel 3.12
zeitliche Gewichtung (Gedächtnis der Leistungsmetrik)	zeitliche Begrenzung des Maximalalters (Zeit nach Abschluss) einer Verkehrsbewegung	abnehmende Gewichtung der Bewegung bei zunehmendem Alter, beschrieben anhand der kumulativen Verteilung der Aktivzeit
Zahlenmenge	kontinuierlich	diskret

Obwohl eine technische Umsetzbarkeit im Versuchsaufbau für beide Metriken möglich erscheint, hat das arithmetische Mittel einen entscheidenden Nachteil: Das arithmetische Mittel ist eine abstrakt Größe. Sie bietet dem Operateur vermutlich keinen vorstellbaren Aufschluss darüber, wie groß der Vorsprung im Wettbewerb zum gegenwärtigen Zeitpunkt ist. Demgegenüber verspricht die Verwendung der Anzahl aktiver Verkehrsbewegungen aufgrund der diskreten Werte eine klare Vorstellung über den Vorsprung zu jedem Zeitpunkt. Aus diesem Grund wird die Anzahl aktiver Verkehrsbewegungen als Leistungsmetrik für den Leistungsvergleich beider Operateure gewählt.

3.2.5.5 Die Vermittlung des Vorsprungs

Der Vorsprung des menschlichen Operators soll nach Leistungsvergleich auf der Arbeitsoberfläche visualisiert und somit vermittelt werden. Der Vorsprung soll zu einem beliebigen Zeitpunkt anhand des Leistungsunterschiedes bestimmt werden, der sich aus der Differenz der Anzahl aktiver Verkehrsbewegungen bildet:

$$n_{\text{Vorsprung}}(t) = n_{\text{aktiv Fluglotsenagent}}(t) - n_{\text{aktiv Mensch}}(t). \quad (3.15)$$

Die technische Lösung zur Bestimmung des Vorsprungs liefert die Schnellzeitsimulation des Fluglotsenagenten anhand eines Vektors, der den zeitlichen Verlauf der Leistungsmetrik des Gesamtszenarios für den Vergleich mit der Echtzeitsimulation enthält. Dieser enthält sortierte Austrittszeitpunkte mit Kennungen des jeweils zugehörigen Luftfahrzeugs und wird während der Schnellzeitsimulation progressiv aufgebaut und in Echtzeit für die visuelle Rückkopplung bereitgestellt. Beide Welten können somit verglichen und der Leistungsunterschied ermittelt werden. Abbildung 35 illustriert exemplarisch den Vorsprung als Differenz zwischen der menschlichen Leistung und der Agentenleistung.

Um dem menschlichen Operator die daraus resultierende verfügbare Zeit zu vermitteln, soll eine korrespondierende Zeitreserve ermittelt werden. Diese Zeitreserve ist jene Zeit, die der Fluglotsenagent zum Aufholen des Vorsprungs benötigt. Sie unterstellt dabei das Einstellen aller Kontrollhandlungen des menschlichen Operators. Dies bedeutet, dass keine Freigaben durch den Menschen erbracht werden. Die Berechnung soll mit Hilfe dieser Annahme und dem vorausberechneten zeitlichen Leistungsverlauf des Fluglotsenagenten vorgenommen werden. Der Leistungsvorsprung wird dabei auf die Zeitachse zu einer Zeitreserve $t_{\text{Vorsprung}}$ abgetragen.

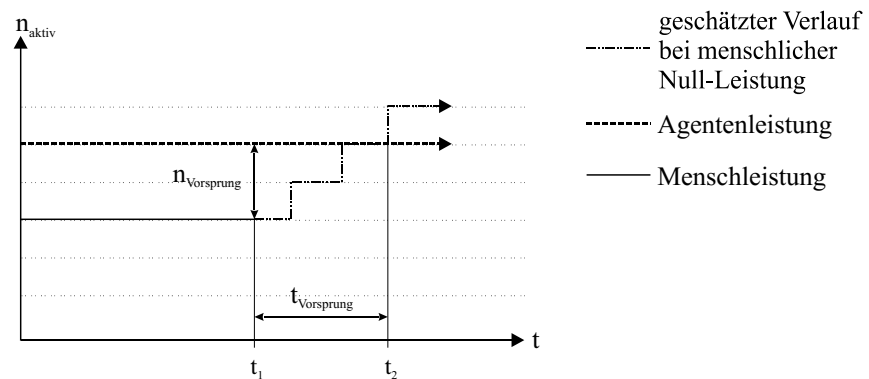


Abbildung 35: Prinzip des Leistungsvergleiches zwischen Fluglotsenagenten und menschlichem Operateur

Der menschliche Operateur wird für eine effektive Vermittlung des Vorsprungs auf zwei Sinneskanälen angesprochen:

- *Visuell*: Die entwickelte visuelle Lösung zeigt die berechnete Zeitreserve als numerischen Wert im Format *mm:ss* an, ergänzt durch eine uhrähnliche Visualisierung (Abbildung 36). Die Uhr ist logarithmisch skaliert, um ein möglichst großes Zeitintervall abbilden und den relevanten Bereich unmittelbar vor Ablauf des Vorsprungs höher auflösen zu können. Unter der numerischen Zeitreserve ist das

Rufzeichen einer Verkehrsbewegung angezeigt, die durch den Fluglotsenagenten als nächstes abgeschlossen wird.

- *Akustisch*: Das akustische Signal spielt Warntöne auf Basis der Zeitreserve in abgestufter Lautstärke und Dringlichkeit zu vordefinierten Punkten auf der Zeitachse. Diese Werte sind 10, 30, 180 und 360 Sekunden.

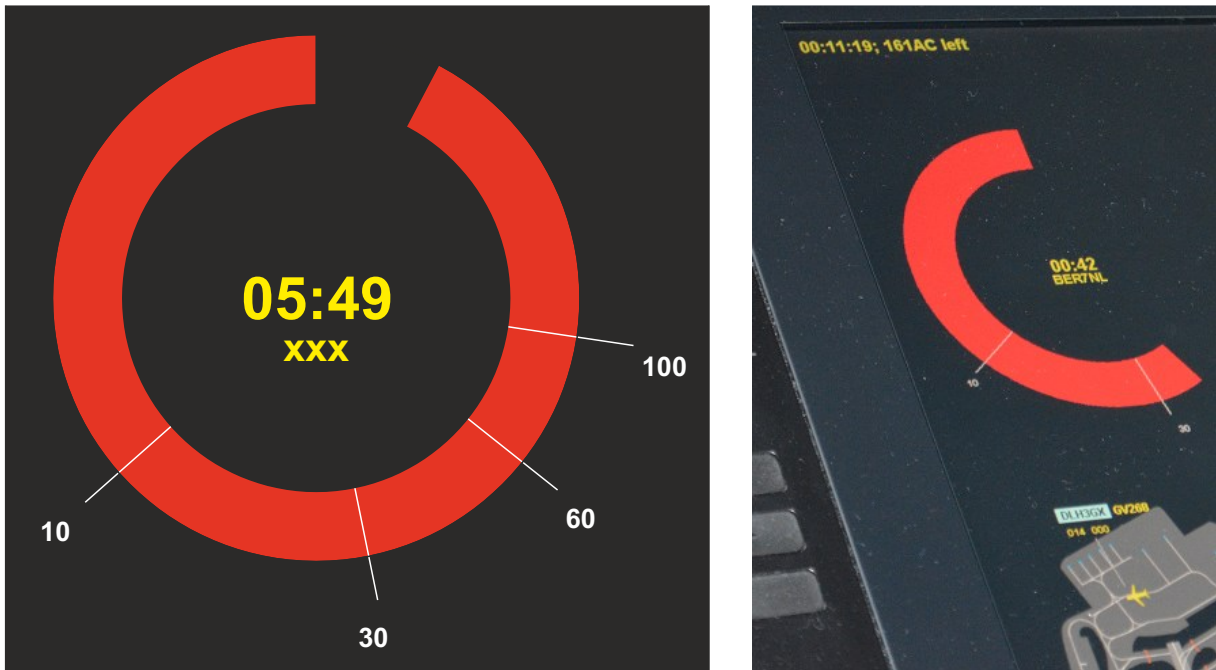


Abbildung 36: Visuelle Anzeige der Zeitreserve des menschlichen Operators als Konzeptskizze (links) und Implementierung (rechts)

3.2.5.6 Das Bestrafungsverfahren

Der Fluglotsenagent hat die Führung in dem Moment übernommen, in dem die Zeitreserve einen Stand von null erreicht. Dieses Ereignis definiert den Zeitfehler ZF . Der Eintritt soll einen Bestrafungsmechanismus aktivieren, der dem menschlichen Operator durch einen unangenehmen Alarmton kenntlich gemacht wird. Die Bedingung für eine derartige Maßnahme ist erfüllt, sobald die Anzahl aktiver Verkehrsbewegungen des menschlichen Operators höher ist als die des Fluglotsenagenten:

$$n_{\text{aktiv Mensch}} > n_{\text{aktiv Fluglotsenagent}} \quad (3.16)$$

Die durch das Szenario gesetzte Anzahl zu kontrollierender Verkehrsbewegungen wird um zwei erhöht und die Arbeitsumgebung des Fluglotsenagenten auf die des menschlichen Operators synchronisiert. Somit werden nach der Bestrafung gleiche Anfangsbedingungen geschaffen, um den Wettbewerb auf einen Gleichstand zurückzusetzen.

3.2.5.7 Das angepasste Zeitbudget

Das Zeitbudget ist ein Maß des objektiven Zeitdrucks. Das hierzu gehörende Konzept wurde bereits in Abschnitt 2.5.3.2 beschrieben. Die Übertragung des Zeitbudgets auf die experimentelle Pilotstudie hat den Vorteil, den induzierten Zeitdruck unabhängig von

Befragungen bestimmen und hinsichtlich der Steuerbarkeit durch die Zeitdrucklast analysieren zu können. Ein wichtiges Maß zur Beobachtung der Effektivität der Zeitdruckinduktion wäre somit geschaffen. Dies setzt allerdings voraus, messbare Größen als Äquivalente der benötigten und verfügbaren Zeit finden zu können. Eine Übersicht möglicher Kandidaten stellt die Tabelle 10 vor.

Tabelle 10: Äquivalenzgrößen zur Definition des angepassten Zeitbudgets

Äquivalent der benötigten Zeit	Äquivalent der verfügbaren Zeit	Zeitbudget
$n_{aktiv\ Mensch}$	$n_{aktiv\ Fluglotsenagent}$	Division
		Differenz ($n_{Vorsprung}$)
	$t_{Vorsprung}$	
t_r	t_L	Division

Aufgrund der konzeptionellen Auslegung des Zeitfehlers wären die folgenden Definitionen

- $n_{aktiv\ Mensch}/n_{aktiv\ Fluglotsenagent}$ oder
- $n_{Vorsprung}$

als geeignete Kandidaten denkbar. Wie in Abschnitt 2.5.2 bereits erläutert wurde, ist der empfundene Zeitdruck das Ergebnis eines subjektiven Risikos für den Zeitfehler. Die Leistungsgrößen sind aufgrund des zeitlosen Bezugs vermutlich wenig geeignet, als Maß des induzierten Zeitdrucks zu dienen.

Als ein geeignetes Maß der verfügbaren Zeit qualifiziert sich demgegenüber die Verwendung der Zeitreserve $t_{Vorsprung}$. Sie ist aufgrund der visuellen und akustischen Rückkopplung maßgeblich für den empfundenen Zeitdruck des menschlichen Operators. Allerdings kann der Zeitreserve konzeptionell keine korrespondierende benötigte Zeit zugeordnet werden.

Alternativ können die Reaktionszeit des menschlichen Operators und die Latenzzeit des Fluglotsenagenten dienen, die gemäß Tabelle 8 korrespondierende Größen der Arbeitsgeschwindigkeiten beider Operateure darstellen. Außer Acht lässt eine solche Definition die Optimierung der Rollweganweisungen und Pistennutzung $Q_{P\&S}$. Hierdurch kann der menschliche Operateur eine höhere Leistung aufweisen und den Vorsprung ausbauen, ohne dass dieser indiziert wird. Aufgrund dessen begrenzt sich die Verwendung des angepassten Zeitbudgets

$$\widehat{TB} = \frac{t_r}{t_L} \tag{3.17}$$

auf eine Indikation des Trends. Sie kann mit dem Zeitfehler ZF und der Leistungsfähigkeit des Probanden korrelieren und in Zusammenhang gebracht werden.

3.2.5.8 Das angepasste Risikomodell

Die Implementierung des Wettbewerbsverfahrens auf Basis des erweiterten Risikomodells (Abbildung 28) und die zugehörigen Annahmen und Definitionen dieses Unterkapitels werden in Abbildung 37 als angepasstes Risikomodell zusammengeführt. Es verdeutlicht den

Vergleich der Leistungsfähigkeit anhand der Anzahl aktiver Verkehrsbewegungen. Der Vergleich wird durch den simulierten Betrieb des Fluglotsenagenten möglich, der eine Referenzleistung unter einer gegebenen Latenzzeit als unabhängige Größe liefert.

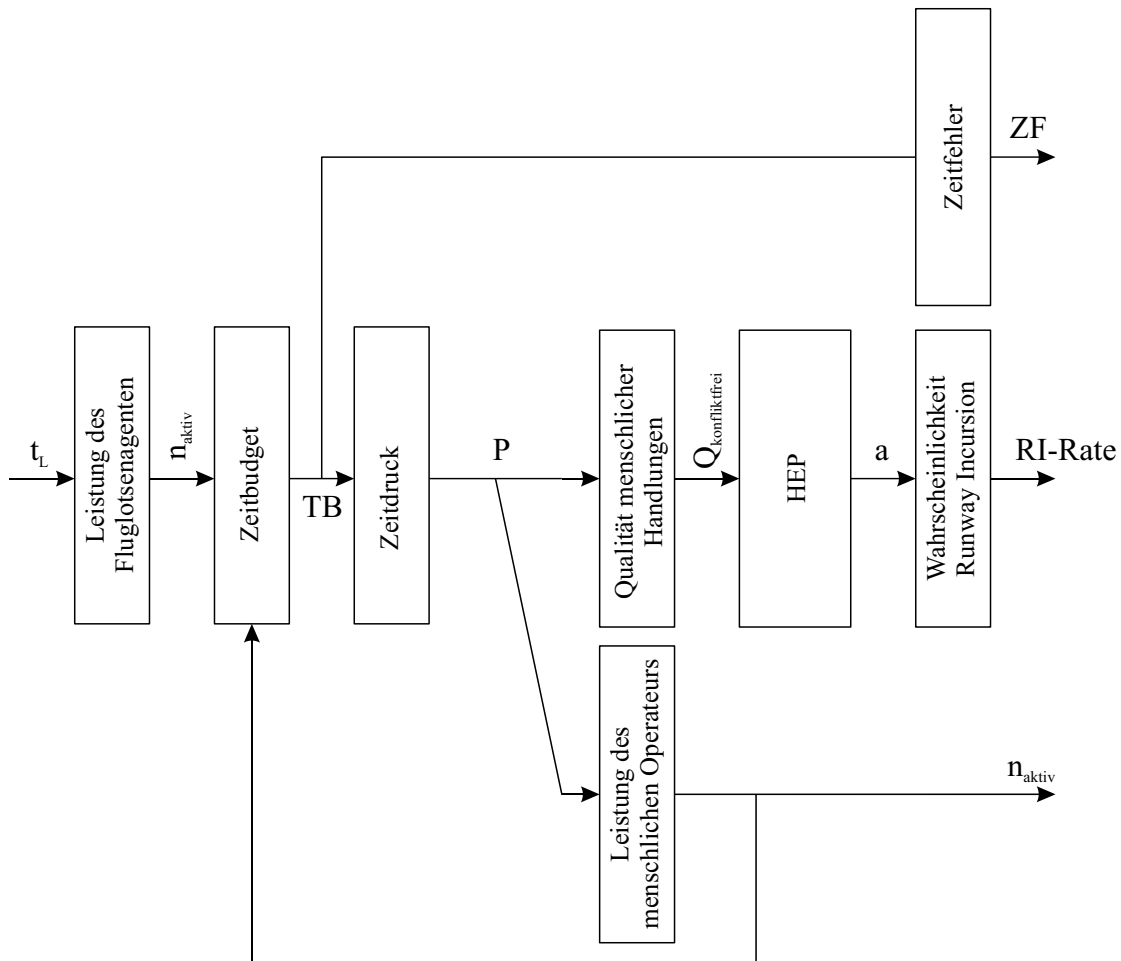


Abbildung 37: Angepasstes Risikomodell

3.2.6 Das Messverfahren

3.2.6.1 Runway Incursion

Für die Bestimmung der Risikometrik stellt die Rollverkehrssimulation einen Algorithmus zur Detektion von Runway Incursion zur Verfügung. Die absolute Häufigkeit dieses Ereignisses, mit Zeitpunkt und Kennung der beteiligten LFZ pro eingetretenem Ereignis, wird für eine Auswertung protokolliert. Das Format ist in Tabelle 11 dargestellt. Die Aufzeichnungsdaten werden dann entsprechend dieses Kopfes als Tabelle in einer Datei abgespeichert.

Tabelle 11: Ausgabeformat des Protokolls – Detektion von Runway Incursion

Echtzeit (hh:mm:ss.sss)	Rufzeichen LFZ A	Rufzeichen LFZ B	Routenname A	Routenname B	Freigabestatus A	Freigabestatus B
----------------------------	---------------------	---------------------	-----------------	-----------------	---------------------	---------------------

Der hierfür zuständige Detektionsalgorithmus überwacht einen definierbaren Abschnitt, der sich in Form eines Rechtecks symmetrisch um die Piste schließt und dabei an den *Runway-*

Holding Position Markings CAT I (ICAO, 2013a) bündig abschließt. Zu jedem Berechnungsschritt der Simulation wird geprüft, ob die Bedingungen für eine Runway Incursion innerhalb dieses Abschnitts erfüllt sind. Diese Bedingungen sind bereits in 3.2.2 erläutert worden. Bei Auswertung des Protokolls können die folgenden Größen ermittelt werden:

- Absolute Häufigkeit der Runway Incursion RI , klassifiziert nach
 - i. Rollend – Landend (TL)
 - ii. Rollend – Startend (TS)
 - iii. Landend – Startend (LS)
 - iv. Startend – Startend (SS)
- Relative Häufigkeit RI -Rate, bei Division der Häufigkeit RI durch die Häufigkeit erfolgter Startvorgänge

3.2.6.2 Benutzeraktionen

Alle Anweisungen und Freigaben, die der menschliche Operateur mit Hilfe der Stiftgeste vornimmt, werden aufgezeichnet. Diese Aktionen werden in einem Format entsprechend Tabelle 12 abgespeichert.

Tabelle 12: Ausgabeformat des Protokolls – Benutzeraktion

Echtzeit (hh:mm:ss.sss)	Simulationszeit	Aktionstyp	Rufzeichen
----------------------------	-----------------	------------	------------

Die abgespeicherten Aktionstypen des Operateurs sind

- absolute Häufigkeit von Freigaben
- absolute Häufigkeit von Rollwegmodifikationen n_{rm}
- Reaktionszeit t_r (Zeitdauer von Anfrage der Freigabe durch ein Luftfahrzeug bis zur zugehörigen Freigabeerteilung)
- Aktionszeit t_{ak} (Zeitdauer der Interaktionsdauer für Freigabeerteilungen und Routenmodifikationen)

3.2.6.3 Flug- und Rollspuraufzeichnung

Die Aufzeichnung der Flug- und Rollspuren erfüllt den Zweck, Reaktionen des Probanden und kritische Zustände des Flughafenverkehrs im Kontext der Bewegungen auf dem Rollfeld und im Anflugbereich erklären zu können. Insbesondere für die Ursachenanalyse oder Berechnung von zeitkritischen Vorgängen können wichtige Beitragsfaktoren des Verkehrsgeschehens ermittelt werden. Das Format ist innerhalb der Datei mit der Endung „obj“ binär abgelegt und kann mit Hilfe von Bibliotheksfunktionen in einem Wiedergabemodus dargestellt werden oder automatisch für die Analyse weiterverarbeitet werden. Zusätzlich wird eine Statistikdatei erstellt, deren Spalten das Format gemäß Tabelle 13 für jede aktive Verkehrsbewegung aufweist.

Tabelle 13: Ausgabeformat des Protokolls – Zusammenfassung Verkehrsbewegungen

Rufzeichen	ICAO Aircraftcode	ICAO Designcode	Richtung (ausgehend/ reingehend)	Strecke [NM]	Aktivzeit [sek]	Wartezeit [sek]	Anzahl Freigaben
------------	----------------------	--------------------	-------------------------------------	-----------------	--------------------	--------------------	---------------------

Folgende Größen können hierdurch ermittelt werden:

- Gesamtmenge der Verkehrsbewegungen
- Häufigkeit gestarteter Verkehrsbewegungen
- Aktivzeiten der Verkehrsbewegungen
- Anzahl der erteilten Freigaben

3.2.6.4 Bestrafungsmaßnahmen

Bei Auslösung des Bestrafungsmechanismus wird ein Eintrag zur Protokollierung angelegt, der dem Format gemäß Tabelle 14 entspricht.

Tabelle 14: Ausgabeformat des Protokolls – Bestrafungsmaßnahme

Echtzeit (hh:mm:ss.sss)	Simulationszeit [sek]
----------------------------	--------------------------

Folgende Größen können hierdurch ermittelt werden: Häufigkeit von Zeitfehlern *ZF*

3.2.7 Die Befragung

Neben der Erfassung der Leistungsfähigkeit des Menschen durch objektive Prozessdaten sollen die empfundene Arbeitsbeanspruchung und der empfundene Zeitdruck mit Hilfe von Befragungen während der Versuchsdurchführung erfasst werden. Dies dient dem Zweck, mögliche Zusammenhänge anhand des Vergleiches von objektiven und subjektiven Daten aus der Zeitdruckinduktion identifizieren zu können. Mögliche Korrelationen könnten Aufschluss darüber bieten, ob das angepasste Zeitbudget ein repräsentatives Maß des vermittelten Zeitdrucks liefert.

Das zur Befragung eingesetzte Befragungsschema lehnt sich an die *Integrated Workload Scale* (IWS) der *University of Nottingham* (Pickup, Wilson, Norris, Mitchell, & Morrisroe, 2005) an und übernimmt die IWS-Skala auf einer 8-stufigen Einteilung (Tabelle 15).

Tabelle 15: 8-stufige Einteilung der Arbeitsbeanspruchung und Zeitdruck für die Befragung

Skala	Arbeitsbeanspruchung/ Zeitdruck
7	Zu hoch
6	Sehr hoch
5	Hoch
4	Akzeptabel
3	Moderat
2	Leicht
1	Minimal
0	Sehr gering

Der Vorteil dieser Einteilung liegt in der Identifikation von Trends und Spitzen bei einem Befragungsintervall von einer, 2, 5 oder 15 Minuten.

3.2.8 Die Szenarientwicklung

Das konzeptionelle Ziel der Szenarien ist, eine hinreichend belastende Arbeitsumgebung für die Messung der Zieleffekte herzustellen. Die wichtigsten Elemente der zu gestaltenden Szenarien orientieren sich an der definierten Primäraufgabe und den Kausalhypothesen. Diese sind

- der Flugplatz mit Pisten, Rollwegen und Parkpositionen,
- die Verkehrsmenge mit dem dazugehörigen Verkehrsmix und
- die Zeitdrucklasten des Zeitdruckinduktionsverfahrens.

Arbeitsbedingungen wie z.B. Wettersituationen, Pistenkontrollen und Betriebsrichtungswechsel werden nicht berücksichtigt.

Für den Versuchsaufbau wird der Betrieb des Flughafens Frankfurt a. Main als repräsentative Arbeitsumgebung gewählt. Innerhalb der Untersuchungen kommt das alte Design von vor dem 20. Oktober 2011 zur Anwendung, dessen drei Pisten in einem *Semi-Mixed Parallel Runway*¹⁷- Betrieb benutzt wurden. Die Entscheidung über eine Startfreigabe hat Abhängigkeiten zu anderen zeitgleichen Start- und Landevorgängen. Die Umgebung bietet somit hinreichend herausfordernde Bedingungen für die Entscheidungsfindung der Probanden. Der simulierte Betrieb umfasst sowohl eingehenden als auch ausgehenden Verkehr, der dem *ICAO Aerodrome Reference Code Letter C*¹⁸ (ICAO, 2013a) entspricht. Zu dieser Klasse Verkehr gehören z.B. die bekannten Flugzeugmodelle der Produktfamilie Airbus A320 und die Boeing 737. Der Verkehr umfasst eingehenden und ausgehenden Verkehr zu gleichen Anteilen. Der landende Verkehr wird hierbei ebenso auf den parallelen Pisten zu gleichen Anteilen verteilt. Die Betriebsrichtung der Pisten bleibt bei allen Szenarien bei 25 bzw. 18.

Tabelle 16: Baseline und Zeitdrucklastszenarien

<i>Szenario</i>	<i>Arbeitsbelastung durch die Verkehrslast</i>	<i>Effektive Stresslast</i>
Baseline	Basislast	Designstress
Zeitdrucklastszenario (ComPerf)	Basislast	Designstress und induzierter Zeitdruck

Für die Pilotstudie werden zwei Klassen von Szenarien entworfen:

- die Baseline und
- die Zeitdrucklastszenarien.

Das Unterscheidungsmerkmal der Zeitdrucklastszenarien gegenüber der Baseline ist die Anwendung des Zeitdruckinduktionsverfahrens. Die Baseline weist sich somit als Szenario mit nicht-modifizierten Arbeitsbedingungen aus und dient als Referenzszenario mit einem Regelbetrieb, der den Designstress definiert. Demgegenüber weist das Zeitdrucklastszenario eine Zeitdrucklast t_L aus, deren Daten einen Zeitdrucklasten-spezifischen Stützpunkt auf der

¹⁷ Die parallelen Pisten bestehen aus einer Landebahn (25L/07R) und einer Piste, die als Start- und Landebahn (25R/07L) benutzt wird. Die dritte Piste (18) hat zwei virtuelle Kreuzungspunkte mit den parallelen Pisten.

¹⁸ Spannweite ist 24m oder größer aber kleiner als 36m/ Äußerer Abstand des Hauptfahrwerks 6m oder größer, aber kleiner als 9m.

Stress-Risiko-Relation herstellen. Die Bedingungen beider Szenarien sind zur Orientierung in Tabelle 16 gegenübergestellt. Die Zeitdrucklastbedingung des Zeitdrucklastszenarios setzt sich aus zwei Zeitdrucklasten zusammen. Dabei überlagert sich der Designstress der Basislast mit dem Stress des induzierten Zeitdrucks. Diese Wechselwirkung wurde in Abschnitt 2.5.4.3 beschrieben. Die Verwendung der Basislast und der damit verbundene zeitliche Verlauf der Verkehrsdichte sind für beide Szenarien gleich.

3.2.8.1 Die Basislast

Die Basislast umfasst eine Verkehrslast, die eine Arbeitsbelastung bei der Ausübung der definierten Kontrolltätigkeiten gewährleisten soll. Die Auslegung des zeitlichen Verlaufs der zu generierenden Verkehrsdichte orientiert sich an zwei Kriterien:

- *Qualitativer Verlauf*: Die Form des Verkehrsprofils soll sowohl Spitzen- als auch Entspannungsphasen umfassen, um die Anpassungen von Freigabestrategien des menschlichen Operateurs mit Hilfe einer variierenden Arbeitsbelastung zu stimulieren. Zu diesem Zweck wird eine Tagesganglinie von Verkehrsbewegungen für den zeitlichen Verlauf der Verkehrsdichte zum Vorbild genommen, die ein typisches Profil internationaler deutscher Flughäfen mit Parallelpistenkonfiguration darstellt. Dieser Tagesgang soll vereinfacht anhand einer Verkehrsdichte nachgebildet werden, die
 - i. drei Spitzenbelastungen (oberes Drittel des maximalen Durchsatzes) als sogenanntes 3-Spitzenprofil umfasst.
 - ii. Das Profil wird durch zwei Entspannungsphasen mit einer geringen Intensität begleitet.

Bei Bestrafung des menschlichen Operateurs soll das Zeitdrucklastenszenario verlängert werden. Es bietet sich hierbei die Wiederholung des Profils an, die eine nahtlose Kontinuität des Verkehrsprofils gewährleistet.

- *Dauer*: Die zeitliche Aufteilung des Szenarios ist in der Reihenfolge wie folgt definiert:
 1. Maximal 90 Minuten für das 3-Spitzenprofil
 2. Maximal 60 Minuten Verlängerung aufgrund der implementierten Bestrafungsfunktion bei Anwenden des Zeitdrucklastszenarios
- *Intensität*: Für die Bemessung der Intensitäten der Amplitudenspitzen und der Entspannungsphasen ist der maximale Durchsatz zu bestimmen. Hierzu werden zwei Grenzen berücksichtigt.
 - i. *Stress-Risiko-Relation*: Die Tagesganglinie soll eine Arbeitsbelastung des Operateurs im mittleren Lastenbereich der Stress-Risiko-Relation gewährleisten, die den menschlichen Operateur gleichzeitig vor Über- als auch Unterforderung schützt.
 - ii. *Erschöpfung der Flughafenkapazität*: Die generierte Verkehrsdichte soll zudem die Rollwege, Pisten und Standplätze des Flughafens zu einem Grad auslasten, bei der die Maximalkapazität der verfahrenskonformen Pisten- und Rollwegebenutzung nicht dauerhaft und systematisch überschritten wird. Das Überschreiten einer derartigen Erschöpfung der Kapazitäten hätte vermutlich Staubildungen an Aufrollpunkten und Rollgassen zur Folge, die die Bedingungen des Wettbewerbs und somit die Beobachtung der Stressreaktionen

des menschlichen Operateurs stören würde. Der Erschöpfungspunkt markiert somit den maximalen Durchsatz für die experimentelle Pilotstudie.

Das konzeptionelle Verfahren zur Kalibrierung der Intensität umfasst eine Verkehrslast, die den Erschöpfungspunkt anhand einer vordefinierten Verkehrsdichte über die Zeit mit abgestuften Intensitäten einzugrenzen und somit zu quantifizieren vermag. Die aus der Verkehrslast resultierende Erschöpfung und die damit verbundenen Staubildungen können anhand der Anzahl aktiver Verkehrsbewegungen n_{aktiv} nachvollzogen werden und bieten Rückschluss auf die Bedingungen, die die Intensität des generierten Verkehrs maximal haben sollte. Das Kalibrierungsverfahren identifiziert den für die Pilotstudie zulässigen Betriebsbereich anhand der folgenden Tests:

- *Dauerstabilität des Betriebes*: Die Stabilität ist allgemein definiert als eine begrenzte Reaktion als Folge einer begrenzten Eingangserregung. Sie bemisst sich bei dem Flugbetrieb auf dem Flugplatz an einer maximal generierbaren Verkehrsdichte, die auf Dauer vorliegen kann. Sie korrespondiert mit der soeben erwähnten maximalen Kapazität der Flughafenressourcen.

Die Quantifizierung dieser maximalen Verkehrsdichte soll mit Hilfe eines Dauerbelastungstests gelingen, der stationäre Verkehrsdichten mit abgestuften Intensitäten anfährt (vgl. Abbildung 38, links). Der zeitliche Verlauf der Verkehrsdichte beschreibt eine gestufte Rampe, die die Intensität stufenartig steigert. Die Entspannungsphasen sollen die Reaktionen der Stufen voneinander entkoppeln und die Unabhängigkeit von der jeweiligen Vorstufe sicherstellen. Die in der Simulation verwendete Szenariokonfiguration ist Anhang C.1 zu entnehmen.

- *Dynamische Belastbarkeit*: Im Gegensatz zur Dauerbelastung können Spitzenlasten eine höhere Amplitude aufweisen. Die Stabilität ist für den Fall nicht verletzt, dass Entspannungsphasen folgen, die in ihrer Dauer den Spitzenlasten angepasst sind. Die maximal mögliche Intensität einer Spitzenbelastung bemisst sich somit an dem Ausgleichsvorgang der Entspannung und den damit verbundenen notwendigen Relaxationszeiten. Die Reaktion soll anhand der Anzahl aktiver Verkehrsbewegungen schließlich nachvollzogen werden.

Der zur Erregung verwendete zeitliche Verlauf der Verkehrsdichte wird nach Vorbild einer Impulserregung ausgelegt. Das dazugehörige Verfahren umfasst die impulsartige Erregung durch eine Aneinanderreihung von Impulsen, die in ihrer Amplitude über n_k stufenartig steigen (vgl. Abbildung 38, rechts).

Die in der Simulation verwendete Szenariokonfiguration ist Anhang C.2 zu entnehmen.

3.2.8.2 Das Zeitdrucklastszenario

Die Implementierung des Zeitinduktionsverfahrens soll in diesem Szenario entsprechend dem Abschnitt 2.5.4 zu Anwendung kommen. Die Parametrierung der Intensität der Zeitdruckinduktion erfolgt dabei anhand der Latenzzeit des Fluglotsenagenten t_l als szenariospezifischer Zeitdrucklastparameter. Zwei Zeitdrucklastparameter sollen zwei Stützpunkte innerhalb der Stress-Risiko-Relation herstellen. Terminologisch dazu passend können die parametrisierten Szenarien als *Laststufen* bezeichnet werden. Für den Nachweis

eines signifikanten Zusammenhangs zwischen unabhängigen und abhängigen Größen dienen die zwei Laststufen mit den Namen

- *ComPerf 1* und
- *ComPerf 2*

mit den zugehörigen Zeitdrucklastparameter t_{L1} und t_{L2} .

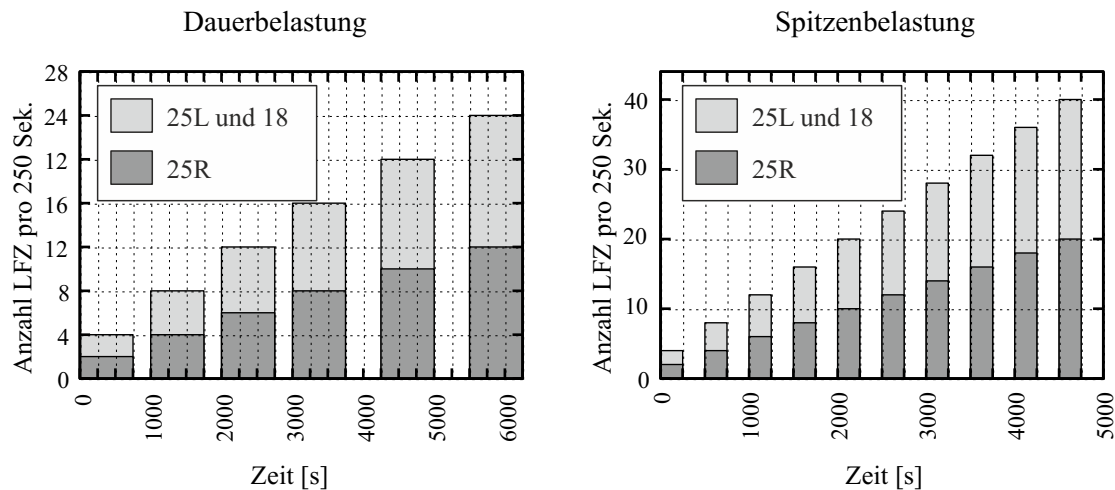


Abbildung 38: Verkehrsbelastungen der Kalibrierungsszenarien Dauerbelastung (links) und Spitzenbelastung (rechts)

Für die Bestimmung geeigneter Zeitdrucklastparameter soll eine Kalibrierung vorgenommen werden. Ziel der Kalibrierung ist es, geeignete Zeitdrucklastparameter in dem oberen Lastenbereich innerhalb der Stress-Risiko-Relation in Abbildung 29, rechts, zu identifizieren. Die Eingrenzung des oberen Lastenbereiches soll mit Hilfe des gemessenen angepassten Zeitbudgets \widehat{TB} , definiert in Formel 3.17, gelingen, welches ein objektives Maß des induzierten Zeitdrucks liefert.

Das Kalibrierungsverfahren induziert dazu eine Zeitdrucklast mit einem zeitlichen Verlauf der Latenzzeit $t_L(t)$. Dabei soll der Verlauf die Extremwerte des Zielintervalls systematisch durchfahren. Das Intervall, in dem die Zeitdrucklastparameter vermutet werden, kann anhand einer vorläufigen Auswertung von Reaktionsdaten am gleichen Versuchsaufbau auf einen Bereich im Intervall $[0 \text{ s}; 150 \text{ s}]$ eingegrenzt werden (Meyer L., Schultz, Schmidt-Roßleben, & Fricke, 2013). Die qualitative Auslegung des Verlaufes orientiert sich an den folgenden Anforderungen:

- *Messung der Stressreaktion*: Identifikation und Quantifizierung eines Zusammenhangs zwischen t_L und den Reaktionszeiten t_r und dem empfundenen Zeitdruck
- *Verzögerungszeit*: Identifikation der Ansprechzeiten des empfundenen Zeitdrucks nach Variation der Zeitdrucklast

Das Kalibrierungsszenario umfasst einen konstanten Verkehrszugang als Basislast. Der zeitliche Verlauf der Reaktionsdaten kann somit auf die Variation der Zeitdrucklast zurückgeführt werden. Der Verkehrszugang ist auf konstante 8,8 Verkehrsbewegungen pro 250 s gesetzt, um die Bedingungen während des Tests gleich zu halten. Diese Verkehrsdichte

hatte sich in vorhergehenden Versuchsdurchführungen mit dem Surface Movement Manager als geeignete Größe erwiesen (Meyer L. , Schultz, Schmidt-Roßleben, & Fricke, 2013).

Das Kalibrierungsverfahren beschreibt zur Erfüllung der soeben erläuterten Anforderungen zunächst einen stufenartigen Verlauf, der die Extremwerte sprunghaft anfährt. Dieser Test soll die maximal hervorrufbare Reaktion des Operateurs hervorrufen und wenn möglich quantifizieren. Die Extremwerte werden dabei nacheinander angefahren und haben eine Intervalldauer von jeweils 900 s, um die Stressreaktion zu erfassen (Abbildung 39). Anschließend beschreibt eine abgestufte Rampenfunktion in 700 s, Schritten eine zunehmende Zeitdrucklast. Dies stellt sicher, den Trend der Stressreaktion verfolgen zu können und mögliche auftretende Effekte mit der dazu vorliegenden Intensität der Zeitdrucklast in Verbindung bringen zu können.

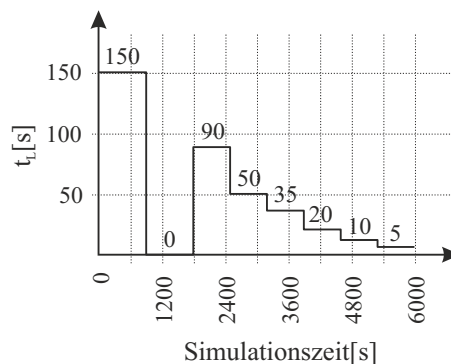


Abbildung 39: Kalibrierungsverlauf der Latenzzeit des Fluglotsenagenten $t_L(t)$ mit Extremwerten zu Beginn [0s; 1800s] und abgestufter Rampe im Anschluss [1800s; 6000s]

Das Befragungsintervall zur Erfassung des empfundenen Zeitdrucks mit Hilfe der IWS-Skala soll 6 Minuten betragen. Diese Zeit wird als hinreichend klein erachtet, um den zeitlichen Verlauf der Reaktion verfolgen zu können. Die in der Simulation verwendete Szenariokonfiguration ist Anhang C.3 zu entnehmen.

3.3 Die Steuerung des Wettbewerbs

Das in Abschnitt 3.2.5.2 entwickelte Leistungsmodell (Abbildung 34) beschreibt eine Abhängigkeit der Aktivzeit und der Leistungsmetrik u.a. von der Entscheidungszeit. Da diese Wirkungskette einen entscheidenden Mechanismus für den Leistungsvergleich der Operateure im Wettbewerb darstellt, ist das korrekte Funktionieren entscheidend für die Steuerbarkeit der Zeitdruckinduktion mit Hilfe der Latenzzeit t_L . Um die Steuerbarkeit der Leistung nach Vorgabe einer vordefinierten Latenzzeit zu überprüfen, soll eine simulationsbasierte Verifikation unter Verwendung des Fluglotsenagenten den erwarteten Effekt einer variierenden Latenzzeit auf die Aktivzeit und die Leistungsmetrik beobachten und die Abhängigkeit bestimmen. Das Leistungsmodell sagt dabei voraus, dass eine kleinere Latenzzeit zu

- einer kleineren mittleren Aktivzeit und in Folge zu
- einer kleineren Anzahl aktiver Verkehrsbewegungen

führt. Als entscheidende Randbedingung der Leistung wird der Einfluss der generierten Verkehrsdichte erachtet, welche für die Verifikation als zusätzlicher Regressor einbezogen wird.

Die Verifikation soll mit Hilfe von Testszenarien durchgeführt werden, die jeweils eingehenden sowie ausgehenden Verkehr anhand einer konstanten Verkehrsdichte symmetrisch auf allen Kanälen generieren. Die Latenzzeit wird über vier Stufen [0 10 20 30] Sekunden und die Verkehrslast über 4 Stufen [33,3 30 26,6 23,3] Sekunden pro Verkehrsbewegung variiert. Die Testszenarien bilden somit 16 Testkonfigurationen. Pro Konfiguration sollen dabei 100 voneinander unabhängige Schnellzeitsimulationen mit je 100 Verkehrsbewegungen gefahren werden. Der Effekt der Latenzzeit auf die beiden Zielgrößen soll mit Hilfe einer multivarianten linearen Regressionsanalyse und dem Bestimmtheitsmaß bestätigt werden.

3.3.1 Die Reaktion der Aktivzeit

Die Vermutung des Zusammenhangs $\hat{t}_{aktiv} = f(t_L)$ wurde bereits in Abschnitt 3.2.5.4 anhand der Formel 3.9 hergeleitet. Für den Fall, dass der angewiesene Rollweg mehrere freigabepflichtige Punkte enthält, hat dies gemäß der Formel einen potenzierenden Effekt der Latenzzeit auf die Aktivzeit. Die Abbildung 40 zeigt rechts beispielhaft die Häufigkeitsverteilung der Aktivzeiten für die zwei Latenzzeiten $t_L = 0\text{ s}$ und $t_L = 30\text{ s}$ bei einem generierten Verkehrszugang von 33,3 Sekunden pro Verkehrsbewegung. Für diese Abbildung wurden aus Gründen der Darstellung 20.000 Stichproben verwendet.

Beide Verteilungen $t_L = 30\text{ s}$ scheinen dabei einer Weibullverteilung nach Vorbild (Günther, Hildebrandt, Fricke, & Strasser, 2011) zu folgen, wenn die Rechtsverschiebung durch die Minimalzeiten $t_{aktiv\ min}$ vernachlässigt wird. Die Form- und Skalenparameter können mit Hilfe der Schätzmethode *Maximum-Likelihood-Methode* geschätzt werden. Bei einer Stichprobenzahl von 10.000 (100 Verkehrsbewegungen bei 100 Testszenarien) führt der Anpassungsgütetest zu signifikanten Testergebnissen. Dies ist auf einen Effekt zu hoher Stichprobenmengen zurückzuführen. Aus diesem Grund wurden die Parameter der Weibullkurven unter Verwendung von 1000 Stichproben ermittelt (Anhang D.1). Die durch den Kolmogorow-Smirnow-Test ermittelten P-Werte schließen die Herkunft der Aktivzeiten aus einer Weibull-verteilten Grundgesamtheit nicht aus.

Abbildung 40, links, illustriert anhand der Boxplots die systematische Verschiebung der Verteilung bei Erhöhung der Latenzzeit. Die Regressionsanalyse (Anhang D.2) belegt die lineare Abhängigkeit der mittleren Aktivzeit zur Latenzzeit und zur generierten Verkehrsmenge pro Zeit mit einem Bestimmtheitsmaß von $R^2 = 94,9\%$ ***. Die Steuerung des Mittels der Aktivzeit durch die Latenzzeit ist somit verifiziert. Anhand der Auswertung konnte die mittlere Anzahl an Freigaben pro Verkehrsbewegung auf 3,17 bei konstantem Verkehrsaufkommen bestimmt werden. Ein möglicher potenzierender Effekt kann allerdings anhand der 4 Stützstellen nicht nachvollzogen werden.

3.3.2 Die Reaktion der Anzahl aktiver Verkehrsbewegungen

Ein möglicher Effekt auf die mittlere Anzahl \hat{n}_{aktiv} kann nach Variation der Latenzzeit beobachtet werden. Die multivariante lineare Regressionsanalyse (Anhang D.3) lieferte in Abhängigkeit der Verkehrslast und der Latenzzeit ein Bestimmtheitsmaß von

($R^2 = 83,8 \%^{***}$) für die gegebenen Konfigurationen. Die Steuerbarkeit der Leistung der Flugplatzkontrolle durch die Latenzzeit gilt somit als verifiziert.

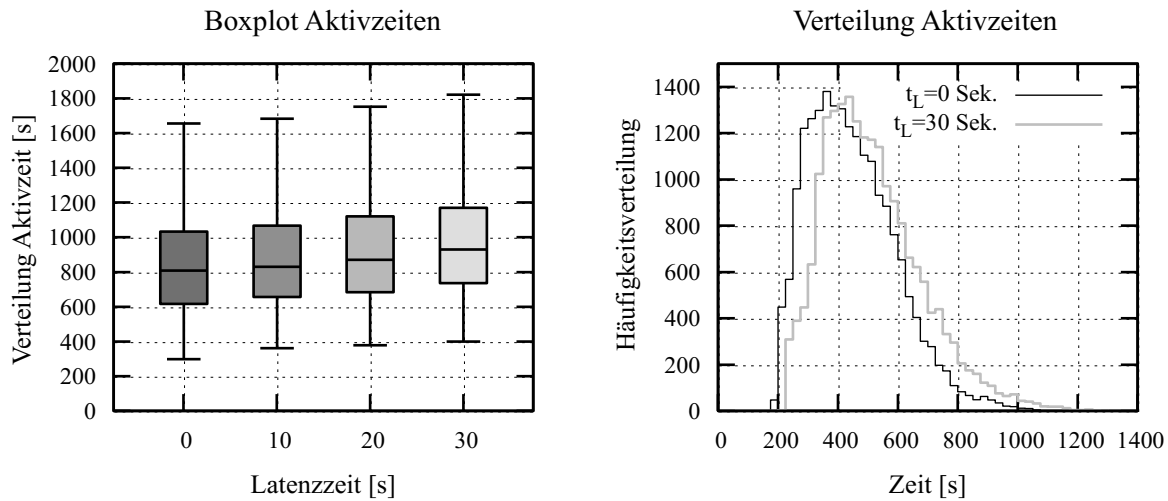


Abbildung 40: Boxplot der Aktivzeitverteilung (links) und Häufigkeitsverteilung der Aktivzeiten (rechts)

Die Boxplots der Abbildung 41 zeigen die Verteilung der Leistungsmetrik am Beispiel von 100 Simulationsläufen über die Simulationszeit des Testszenarios. Das stufenartige Einsetzen bei $t = 0$ s und Aussetzen bei $t = 3333$ s des generierten Verkehrs führt hier zu einer Sprungantwort der Leistungsmetrik. Auf Basis des zeitlichen Verlaufes des Boxplots kann der Relaxationsvorgang beobachtet werden, bei der der Ausgleichsvorgang zwischen Verkehrszugang- und abgang eine Konvergenz von n_{aktiv} gegen einen Grenzwert bewirkt. Dieses Verhalten kann anhand des Leistungsmodells, Abbildung 33, erklärt werden. Auf Grundlage der in Abbildung 41 vorgestellten Verläufe können zwei Relaxationszeiten¹⁹ bestimmt werden. Sie liegen je nach Latenzzeit zwischen 4 und 5 min und sind ein Maß der systembedingten Verzögerung der Leistungsmetrik (Tabelle 17). Bei Start der Simulation und bis zur Herstellung der stationären Lastbedingungen sollten aufgrund dieser Verzögerung bei Latenzzeiten die kleiner als 30 s sind ca. 5 min eingeplant werden. Bei größeren Latenzzeiten wären die Relaxationszeiten durch zusätzliche Simulationsergebnisse zu überprüfen.

Tabelle 17: Relaxationszeiten

t_L [s]	Relaxationszeiten [s]
0	254,6
30	296,3

3.3.3 Das Verifikationsergebnis

Sowohl die Transferfunktionen, aufgeführt in Anhang D.2 und Anhang D.3, als auch die ermittelten Relaxationszeiten geben Aufschluss über den zu erwartenden Verlauf der

¹⁹ Die Relaxationszeit τ ist eine spezifische Zeitkonstante eines zeitdynamischen Relaxationsprozesses mit exponentiellen Relaxationsverhalten. Anhand der Form $e^{-\frac{t}{\tau}}$ definiert die Zeit den Punkt, an dem ca. 63,2% des Relaxationsprozesses erreicht sind.

Leistungsmetrik des Fluglotsenagenten bei gegebener Verkehrslast und Latenzzeit. Das Leistungsmodell, Abbildung 34, kann anhand einer derartigen Verifikation nachvollzogen werden und den untersuchten Größen eine Steuerbarkeit durch die Latenzzeit bescheinigt werden. Der Wettbewerb ist somit mit Hilfe der Latenzzeit des Fluglotsenagenten steuerbar.

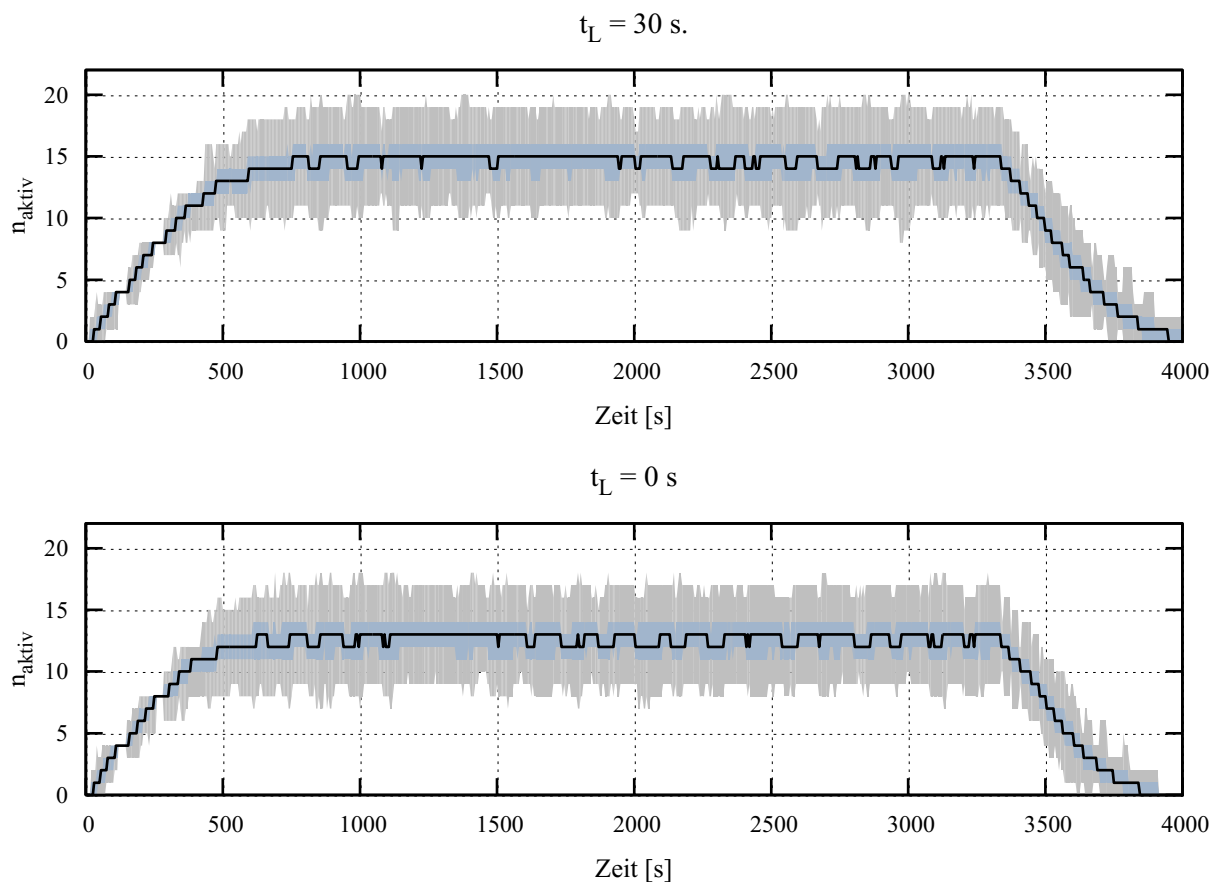


Abbildung 41: Anzahl der aktiven Verkehrsbewegungen in der Simulation über die Zeit, mit den Boxplot-Werten Median, $Q_{0,25}$, $Q_{0,75}$, $Q_{0,1}$, $Q_{0,9}$, bei Variation der Latenzzeit t_L

3.4 Die Organisation der Durchführung

3.4.1 Die Akquisition der Probanden

Für die Auswahl geeigneter Probanden wurden entsprechend der Anforderung des Abschnitts 3.1.3.4 14 studentische Kandidaten aus dem Studiengang Verkehrsingenieurwesen des 5. Semesters der TU Dresden akquiriert. Sie wurden im Dezember 2013 jeweils mit Hilfe von Trainingsszenarien zu zwei Terminen je 1,5 Stunden geschult und mit den Aufgaben des Platzverkehrslotsen, beschrieben in Abschnitt 3.2.3, sowie der Handhabung der Versuchsumgebung vertraut gemacht. Es wurden zusätzlich Bewertungsbögen angefertigt, die sowohl die

- Selbstorganisation der Kandidaten, wie z.B. das Aneignen von notwendigem Fachwissen vorab der Durchführung und Pünktlichkeit, als auch
- das Geschick und die Lernfähigkeit im simulierten Betriebsablauf als Maß der persönlichen Leistungsfähigkeit

dokumentierten. Das Muster des dazu entwickelten Kandidaten-Bewertungsbogens ist in Anhang B.1 hinterlegt. Für die Auswahl geeigneter Kandidaten wurde ein Ranking erstellt, das alle Kandidaten entsprechend ihrer qualitativen Leistungseinschätzung bewertete. Die Bewertungen jedes Kandidaten wurden in einem Paar-Vergleich relativ zu den anderen Kandidaten bestimmt und eine entsprechende Position innerhalb des Rankings zugeordnet. Entsprechend der Reihenfolge dieses Rankings wurden die ersten drei Kandidaten für die Fortsetzung der Teilnahme als Probanden eingeladen. Die Bezeichnungen lauten im Folgenden Proband A, B und C.

3.4.2 Die Versuchsplanung

Die Planung der Versuchsdurchführung umfasste sowohl die Kalibrierung als auch den Hauptversuch. Die drei Versuchskonfigurationen der Hauptversuche bestanden aus

1. der Baseline und
2. den Zeitdrucklastszenarien ComPerf 1 und 2.

Die Reihenfolge der drei Konfigurationen wurde für die Kompensation von Lerneffekten je Kandidat und Runde systematisch permutiert. Die gewonnenen Ergebnisse sollten auf diese Weise eine Unabhängigkeit von der Reihenfolge der Szenarien erlangen und Sequenzeffekte vermieden werden (Bortz & Döring, 2006). Im Vorfeld der Hauptversuche wurden die Kalibrierungstests der Basislast als auch der Zeitdrucklastszenarien durchgeführt. Für eine Bestimmung des Einflusses des Lerneffektes waren die drei Versuchskonfigurationen je Proband in zwei Durchführungen aufgeteilt. Die vollständige Übersicht aller Probanden ist in Anhang F zusammengefasst.

Der Versuchszeitraum umfasste die Werktage von 4 Wochen, bei dem jeder Tag in drei Zeitfenster (vormittags, nachmittags und abends) eingeteilt wurde. Die Versuchstermine wurden bezüglich der zugeteilten Tageszeit ebenso systematisch variiert, um tagesformbedingte Einflüsse auszuschließen zu können. Die Versuchsserie und die Zusammensetzung der Konfigurationen sind zum Zwecke der Übersicht in Tabelle 18 zusammengefasst.

3.5 Die Kalibrierung der Basislast

Die Kalibrierung der Basislast wurden entsprechend der in Abschnitt 3.2.8 beschriebenen Tests durchgeführt. Der eingehende und ausgehende Verkehr wurde dazu nach der vordefinierten Verkehrsdichte in Abbildung 38 generiert und die Verkehrslast hergestellt. Der durch die Probanden A, B und C kontrollierte Flugbetrieb hatte einen zeitlichen Verlauf der Anzahl aktiver Verkehrsbewegungen n_{aktiv} , der in Abbildung 42 für beide Tests wiedergegeben wird. Zu erkennen ist die operatorspezifische Leistungsfähigkeit, die Verkehrsbewegungen entsprechend der Verfahren zur Pistenstaffelung und Rollwegnutzung optimal zu kontrollieren.

Die drei Verläufe des Szenarios Dauerbelastung deuten auf die Existenz eines Grenzwerts für die Intervalle 1 und 2 (bis 1750 s) hin, gegen den der Verlauf von n_{aktiv} für alle Probanden zu konvergieren scheint. Ab Intervall 4 (3000 bis 4000 s) kann eine solche Konvergenz für keinen Probanden festgestellt werden. Der divergierende Trend der Anzahl aktiver Verkehrsbewegungen wird lediglich aufgrund der einsetzenden Entspannungsphase auf eine Spitze von 35 bis 40 Verkehrsbewegungen begrenzt. Dies ist auf den

Erschöpfungseffekt der Flughafenkapazitäten zurückzuführen, der einen Verlauf ohne Ausgleichsprozess hervorruft. In diesem konkreten Fall handelt es sich um die Pistenbelegung, die aufgrund des landenden Verkehrs über keine freien Kapazitäten für startenden Verkehr verfügt. Der zeitliche Verlauf des Probanden A deutet dabei bereits in Intervall 3 (2000 bis 3000 s) auf eine solche Erschöpfung hin. Es kann geschlussfolgert werden, dass dieser Effekt allgemein zwischen Intervall 3 und 4 einsetzt. Schlussfolgernd sollte die Basislast eine Verkehrsdichte von 12 pro 250 s nicht im Mittel überschreiten. Um auch bei Spitzenlasten und schwankenden Leistungen des Platzverkehrslotsen eine hinreichende Sicherheit gegenüber Staubildungen aufzuweisen, wird eine mittlere Verkehrsdichte von 8 pro 250 s (66 %ige Last) mit 4 Verkehrsbewegungen pro 250 s Toleranz gegenüber dem Maximum als hinreichende mittlere Basislast erachtet.

Tabelle 18: Versuchsconfigurationen pro Proband

<i>Versuchsconfiguration</i>		<i>Durchführungen pro Proband</i>	<i>Gesamtzeit pro Proband [h]</i>	
Training und Qualifikation der Kandidaten		2	3	
Kalibrierungstests zur Parametrierung der Tagesganglinie der Basislast		2	3	
Kalibrierungstests zur Parametrierung der Latenzzeiten der Laststufen ComPerf 1 und 2		1	2	
Hauptversuche	Baseline	2	4	
	ComPerf	Laststufe 1	2	4
		Laststufe 2	2	4
Gesamt		11	20	

Das Ergebnis des Szenarios Spitzenbelastung zeigt die Leistungsfähigkeit des Operators, den Flugverkehr unter impulsartigen Belastungen effektiv anweisen zu können. Die zeitliche Entwicklung der Anzahl an Verkehrsbewegungen in Abbildung 42 zeigt eine kontinuierlich steigende Menge aktiver Verkehrsbewegungen, die die Kapazität des Flughafens ebenso wie bei der Dauerbelastung zunehmend belegen und erschöpfen. So treten insbesondere Staus an den Aufrollbereichen zur Piste auf, die während des einsetzenden Impulses durch den ebenfalls intensivierten Landebetrieb zunehmend belegt ist. Ab Intervall 5 (2000 bis 2250 s) setzt eine Erschöpfung der Kapazität ein, die den kontinuierlichen Fluss ausgehender Verkehrsbewegungen über die Pisten aufgrund der Grundlast nicht mehr gewährleisten kann. Dies kann anhand der abgeflachten Spitzen der Intervalle 5, 6 und 7 in Abbildung 42 nachvollzogen werden, bei der die abfallende Flanke in zwei Abschnitte mit unterschiedlichen Abfallraten unterteilt werden kann. Details dazu sind in (Kühn, 2014, S. 51) beschrieben. Der Effekt tritt ab einer generierten Verkehrsdichte von 16 Verkehrsbewegungen pro 250 Sekunden auf. Um den simulierten Flugbetrieb vor Erschöpfungseffekten der Flughafenkapazität zu schützen, soll die Spitzenlast der Tagesganglinie aus diesem Grund eine derartige Verkehrsdichte nicht überschreiten.

Für die Auslegung einer Tagesganglinie werden die folgenden Merkmale mit den Anforderungen der Szenarioentwicklung 3.2.8.1 abgestimmt und festgelegt:

Die Kalibrierung der Basislast

- *Anzahl der Verkehrsbewegungen pro Szenario:* Die Dauer der Tagesganglinie wird für eine optimale Konfiguration der Flugplandatei des Rollverkehrssimulators auf 5000 s festgelegt, die das geforderte Maximum von 90 min erfüllt. Aus einer mittleren Verkehrsdichte von 8 Bewegungen pro 250 s resultiert eine zu generierende Anzahl von 160 Bewegungen.
- *Verteilung der Verkehrsbewegung:* Die Anzahl wird symmetrisch nach startendem und landendem Verkehr unterschieden und zu gleichen Teilen über die je zwei Anflug- und Abflugkanäle der Simulation verteilt.

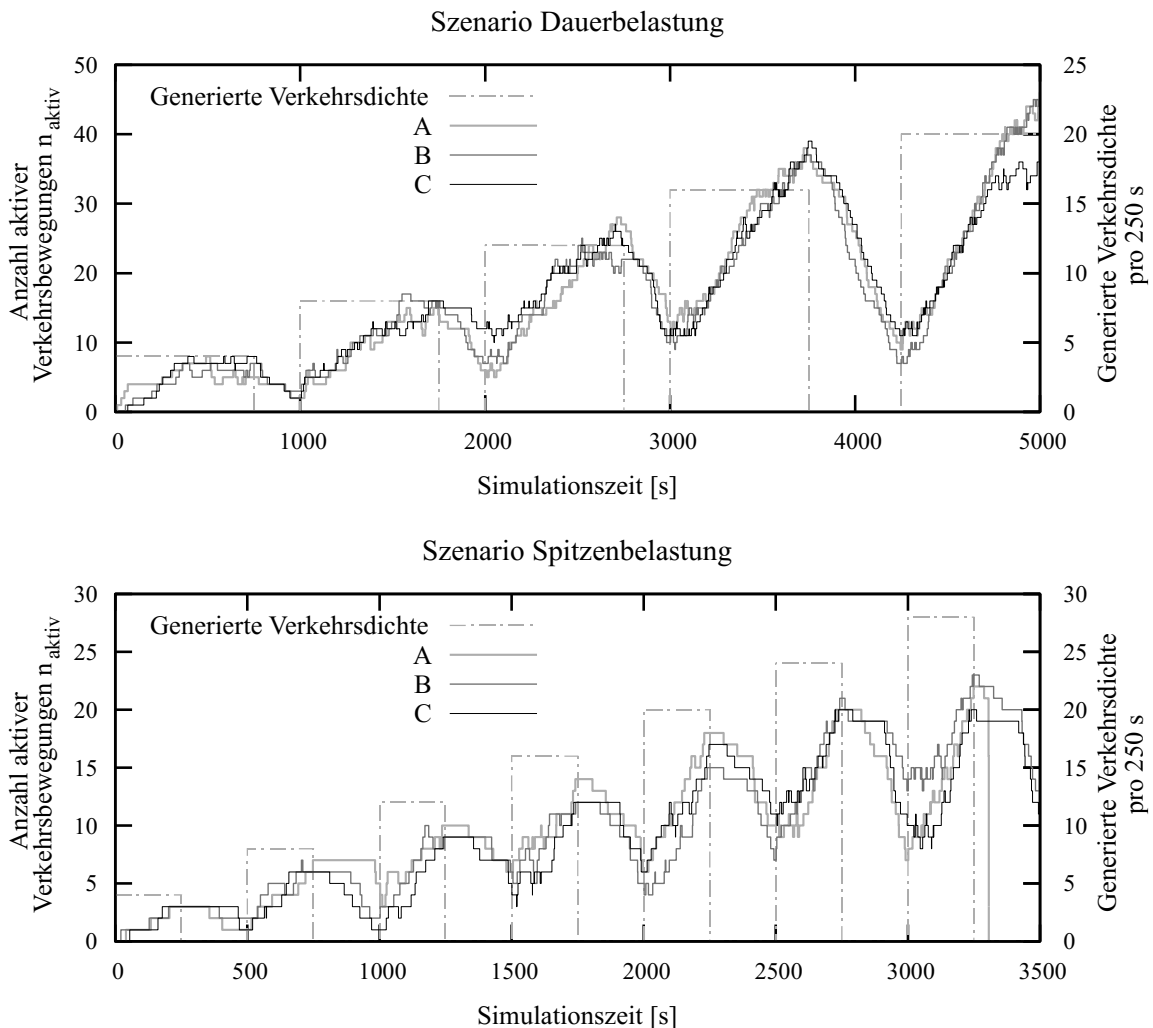


Abbildung 42: Zeitlicher Verlauf der Anzahl aktiver Verkehrsbewegungen des Kalibrierungsszenarios der Basislast

- *Anzahl der Entscheidungssituationen:* Die gewählte Anzahl von Verkehrsbewegungen führt gemäß dem Ergebnis der Leistungsverifikation in Abschnitt 3.3.1 im Schnitt zu 507 Entscheidungssituationen über zu erteilende Freigaben. Bei zwei Runden (vgl. Versuchsplanung 3.4.2) entspricht dies ca. 1014 Entscheidungssituationen pro Proband und Laststufe bzw. Baseline. Durch die symmetrische Proportion von startendem und landendem Verkehr können pro Versuchsdurchführung mindestens 80 Startfreigaben erwartet werden, die zu einer Runway Incursion führen können. Dies entspricht pro Lastenstufe und Baseline á zwei Runden mindestens 160 sicherheitsrelevanten

Startfreigaben, die somit die geforderte Stichprobenanzahl von wenigstens 100 ebenso erfüllen. Nicht eingerechnet sind die Rollfreigaben zur Kreuzung von sicherheitsrelevanten Verfahrensbereichen, wie z.B. die Piste oder der Anflugbereich auf den Rollwegen Sierra und Romeo. Ebenso wird hierbei die ebenfalls als sicherheitsrelevant zu erachtende Aufrollfreigabe vernachlässigt, die aber in enger Verbindung zur Startfreigabe erteilt wird. Aufgrund der paarweisen Erteilung beider Freigaben wird die Aufrollfreigabe nicht als eine unabhängige Gelegenheit für eine Runway Incursion erachtet.

- *Verlängerung der Tagesganglinie aufgrund von Bestrafungen:* Bei Anwendung der Baseline innerhalb der Zeitdrucklastszenarien mit den Laststufen ComPerf 1 und 2 kann die Anwendung des Bestrafungsverfahrens dazu führen, dass die effektive Anzahl an Verkehrsbewegungen um die doppelte Anzahl an Bestrafungen höher ausfallen kann. In diesem Fall verlängert sich das Szenario um die entsprechende Anzahl, in dem die Tagesganglinie wiederholt wird.

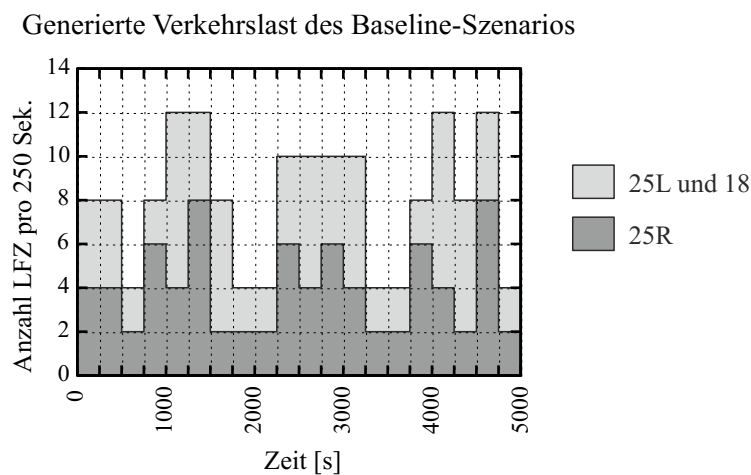


Abbildung 43: Tagesganglinie der generierten Verkehrsmenge der Basislast mit drei Spitzendurchgängen

Auf Grundlage der zusammengetragenen Merkmale wurde ein zeitlicher Verlauf der Verkehrsdichte entworfen (Abbildung 43). Die Tagesganglinie der Basislast wurde zur Dimensionierung der Mengen in zwei Gruppen aufgeteilt:

- 25R (Start- und Landebahn)
- 25L/ 18 (Landebahn/ Startbahn)

Die zwischen den Spitzen eingeplanten Entspannungsphasen begrenzen mit einem Amplitudenminimum von ca. 60 Sekunden pro Verkehrsbewegung (4 LFZ pro 250 Sekunden) die eingehende Verkehrsdichte nach unten.

3.6 Die Kalibrierung der Zeitdrucklastszenarien

Zur Identifikation geeigneter Zeitdrucklastparameter t_{L1} und t_{L2} wurde das Kalibrierungsverfahren angewendet. Nach Anfahren der Laststufen wurden

- das Zeitbudget anhand der aufgezeichneten Reaktionszeiten für die Entscheidungsfindung (t_r),

- die Befragungsdaten zu Arbeitsbeanspruchung und Zeitdruck nach IWS-Skala,
- die aufgezeichnete Anzahl Runway Incursions und
- die aufgezeichnete Anzahl an Zeitfehlern

gemessen und analysiert. Für die Analyse wurde der Verlauf aller Messgrößen und gesetzte Last t_L in einer zeitkontinuierlichen Darstellung dargestellt (Abbildung 44). Die Abbildung zeigt alle Probanden im Vergleich auf 6200 s Versuchszeit mit den folgenden drei Kurven je Proband: Kalibrierungstrajektorie t_L (schwarz), Reaktionszeit t_r (gestrichelt schwarz) und das Zeitbudget (grau). Durch Bestrafung verursachte längere Versuchszeiten wurden zugunsten der Vergleichbarkeit nicht in das Diagramm übernommen. Auf den Verlauf der Reaktionszeit t_r wurde ein Glättungsverfahren angewandt, das mit Hilfe eines gleitenden Mittels, mit beidseitiger symmetrisch-linearer Gewichtung und 200 s Fensterbreite realisiert wurde. Die zeitdiskreten Reaktionszeiten konnte anhand eines derartigen Verfahrens zu einer zeitkontinuierlichen Kurve abgetragen werden. Mögliche Tendenzen des Verlaufes der Arbeitsgeschwindigkeit können so nachvollzogen und mit der korrespondierenden Latenzzeit in Verbindung gebracht werden.

3.6.1 Die Analyse der Reaktionszeit

Der Zeitverlauf der Reaktionszeit folgt einem zyklischen Verhalten mit einer ungefähren Periodendauer von 270 s für allen Probanden (Abbildung 44). Ein derart oszillierender Verlauf könnte durch das Arbeitsmuster der damit verbundenen Verteilung der Aufmerksamkeit erklärt werden. So könnte der Bereich der Aufmerksamkeit zwischen dem sicherheitsrelevanten Pistenbereich und den Parkpositionen zu einer Schwankung in der Reaktionszeit führen. Hohe Reaktionszeiten können somit durch lange Wartezeiten der Verkehrsbewegungen auf der Parkposition entstehen, während der Proband seine Aufmerksamkeit auf die Piste richtet. Die Spitzen im Verlauf der Reaktionszeit können also darauf hindeuten, dass der Proband die Aufmerksamkeit auf die wartenden Verkehrsbewegungen an den Parkpositionen richtet.

Entgegen den Erwartungen ist die Arbeitsgeschwindigkeit t_r nicht durch die ändernden Zeitdrucklasten zu beeinflussen. Dies bestätigt die lineare Regressionsanalyse, deren Ergebnis keinen Schluss einer Abhängigkeit zulässt (P-Wert > 5 %). Die Durchführung der Regressionsanalyse ist in Anhang H.1 detailliert vorgestellt. Es ist anzunehmen, dass die Effektstärke bei ca. 600 Stichproben der Reaktionszeiten zu gering ausfällt. Ebenso könnte es sich um einen Zusammenhang im Modell handeln, bei dem die Reaktionszeit des Menschen keine Änderung in der Grundgesamtheit erfährt und die Variation der Zeitdrucklast ohne Auswirkungen ist.

Sollte die Reaktionszeit keinen Effekt durch die Zeitdrucklast erfahren, so wäre das angepasste Zeitbudget eine antiproportionale Funktion der Latenzzeit $TB \sim 1/t_L$. Die lineare Regression ist Anhang H.2 zu entnehmen und bestätigt die Vermutung mit einem signifikanten*** Zusammenhang, gezeigt in Tabelle 19. Bei Proband C fällt das niedrige Bestimmtheitsmaß auf, bei dem die Reaktionszeit für eine höhere Varianz sorgt.

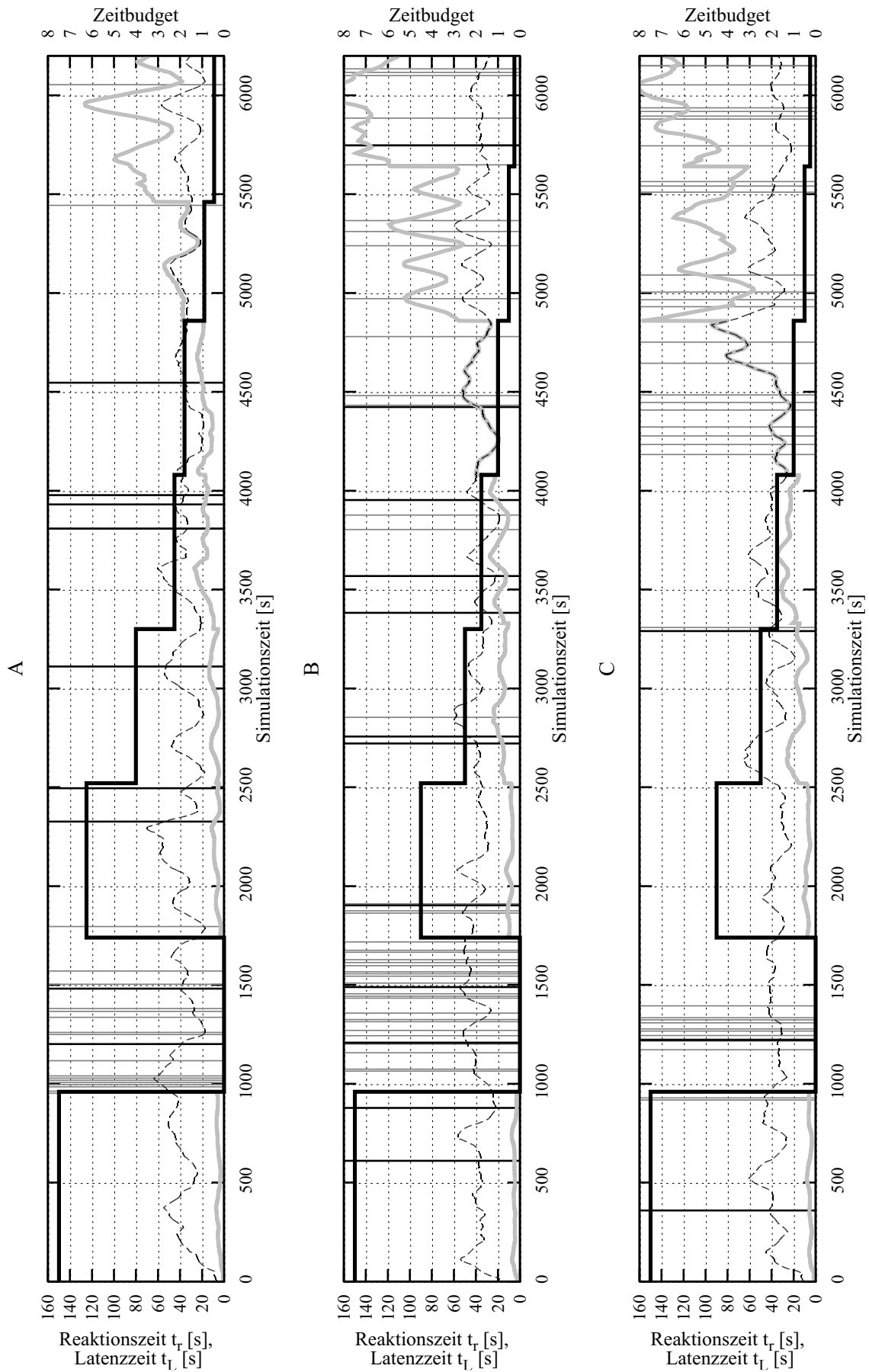


Abbildung 44: Zeitverlauf der Reaktionszeit t_r (gestrichelt), Latenzzeit t_l (schwarz) und das Zeitbudget (grau) der drei Probanden A, B und C mit den Ereigniszeitpunkten der Runway Incursion (schwarz) und Zeitfehlern (grau)

3.6.2 Die Analyse des Zeitfehlers

Deutlich zu erkennen sind in der Abbildung 44 Bereiche, die eine Verdichtung der Zeitfehler aufweisen. Dies gilt insbesondere für die Intervalle mit $t_L \leq 20s$. Dies bestätigt die Abbildung 45 links, die bei abnehmender Latenzzeit einen steigenden Trend der Häufigkeiten der Zeitfehler verzeichnet. Die Prüfung anhand einer Regressionsanalyse bestätigt den signifikanten*** Zusammenhang $ZF \sim 1/t_L$ für alle Kandidaten in Tabelle 19 und liefert ein Bestimmtheitsmaß, das auf einen überwiegend linearen Trend schließen lässt. Die dazugehörige Berechnung ist in Anhang H.3 hinterlegt.

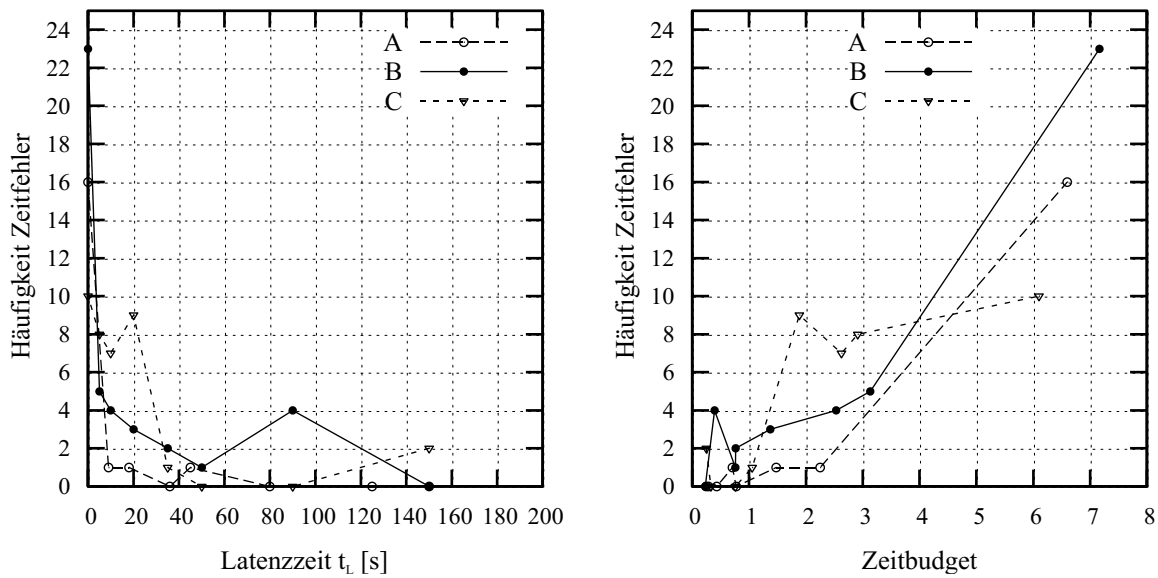


Abbildung 45: Häufigkeiten der Zeitfehler der Probanden über die gesetzte Latenzzeit t_L (links) und über das mittlere Zeitbudget (rechts)

Die Abbildung 45 rechts zeigt die Häufigkeit an Zeitfehlern über das mittlere Zeitbudget. Auffällig ist hierbei der Anstieg nach der Anzahl an Zeitfehlern nach Überschreiten des Schwellenwertes $\widehat{TB} > 1$. Der Zusammenhang $ZF \sim \widehat{TB}$ lässt sich mit Hilfe einer Regressionsanalyse prüfen. Das Ergebnis liefert einen signifikant*** linearen Zusammenhang für alle Kandidaten (Tabelle 19). Die dazugehörigen Berechnungen sind Anhang H.4 zu entnehmen.

Das „Abrutschen“ des Probanden in den Bereich $TB > 1$ ist ein Anzeichen dafür, dass der Proband im Wettbewerb nicht bestehen kann. Da er oder sie den Vorsprung gegenüber dem Fluglotsenagenten nicht ausbauen kann, ist es zu vermuten, dass der Betrieb innerhalb dieses Bereiches erschöpfend auf den Operator wirkt. Zusätzlich könnte dieser Arbeitsbereich zu einer zunehmenden Unempfindlichkeit des Menschen gegenüber der Bestrafung führen, da der Wettbewerb ausweglos außerhalb seines Leistungsbereiches zu liegen scheint. Die Grenze zur Volllast bei $\widehat{TB} = 1$ markiert aus diesem Grund den maximal ausschöpfbaren Bereich des Induktionsverfahrens, in dem der Proband operieren sollte.

Der für die Pilotstudie interessante obere Lastbereich kann aus diesem Grund bei einer Latenzzeit zwischen 20 und 40 s pro Proband eingegrenzt werden. Der Bereich ist anhand

der Abbildung 44 nachvollziehbar, die den Bereich für die geschätzten Latenzzeiten unmittelbar unterhalb eines Zeitbudgets von 1 verortet.

Tabelle 19: Regressionsanalyse des Zusammenhangs zwischen Häufigkeit der Zeitfehler und mittlerem Zeitbudget

Proband	Bestimmtheitsmaß R^2		
	t_L^{-1} zu \widehat{TB}	t_L^{-1} zu ZF	\widehat{TB} zu ZF
A	95,1***	99,4***	96,6***
B	92,1***	97,5***	92,7***
C	86,3***	37,9***	61,9*

3.6.3 Die Analyse der subjektiven Arbeitsbeanspruchung und des Zeitdrucks

Die empfundene Arbeitsbeanspruchung und der Zeitdruck konnte mit Hilfe der Befragung auf der 8-stufigen Skala erfasst werden. Der Werteverlauf über die Zeit (Abbildung 46) zeigt sowohl bei der Arbeitsbeanspruchung als auch bei dem Zeitdruck eine Reaktion auf die Latenzzeit, die eine Abhängigkeit darstellen könnte. Der Werteverlauf des Zeitdrucks scheint einen stärkeren Effekt durch die Zeitdrucklast zu erfahren als die Arbeitsbeanspruchung, da der Werteverlauf des Zeitdrucks der gesetzten Latenzzeit t_L getreuer folgt.

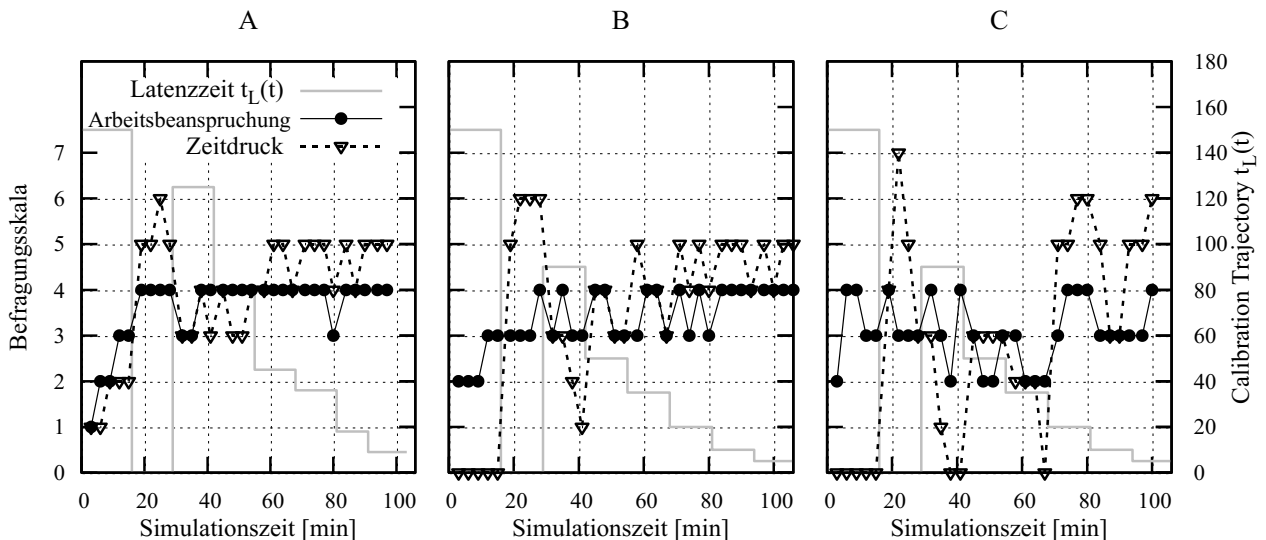


Abbildung 46: Verlauf der Arbeitsbelastung und des Zeitdrucks für die Probanden über die Versuchszeit

Die lineare Regression bestätigt die Relation zwischen der Latenzzeit t_L und dem Zeitdruck. Die Tabelle 20 zeigt hierzu den Unterschied der Reaktionen. Der Zeitdruck weist aufgrund der Bestimmtheitsmaße eine größere Empfindlichkeit auf. Am deutlichsten fällt dieser Unterschied bei Kandidat C aus, der keinen Zusammenhang bei der Beanspruchung erkennen lässt. Demgegenüber können 63 % der Varianz des Zeitdrucks anhand der Latenzzeit erklärt werden. Die Zeitdruckwerte der Kandidaten A und B lassen ebenfalls einen signifikant*** linearen Zusammenhang erkennen ($p < 10^{-3}$). Das XY-Diagramm Abbildung 47 zeigt den Unterschied zwischen den beiden Befragungen über das mittlere Zeitbudget des jeweiligen

Lastintervalls. Die Berechnungen hierfür sind dem Anhang H.5 und Anhang H.6 zu entnehmen.

Tabelle 20: Bestimmtheitsmaße der linearen Regression von t_L zu Arbeitsbeanspruchung und Zeitdruck

Proband	Bestimmtheitsmaß R^2 [%]	
	Arbeitsbeanspruchung	Zeitdruck
A	48,7***	81,4***
B	43,3***	90,7***
C	0,0	63,9***

Es prägt sich statt eines linearen Zusammenhangs eine Kurve aus, die einem asymptotischen Verlauf ähnelt und sich anscheinend einem Grenzwert nähert (Abbildung 47 rechts). Dieser Verlauf ist insbesondere bei den Kandidaten A und B deutlich zu erkennen, Kandidat C lässt den Verlauf errahnen, weist jedoch deutliche Abweichungen gegenüber einem asymptotischen Verlauf auf.

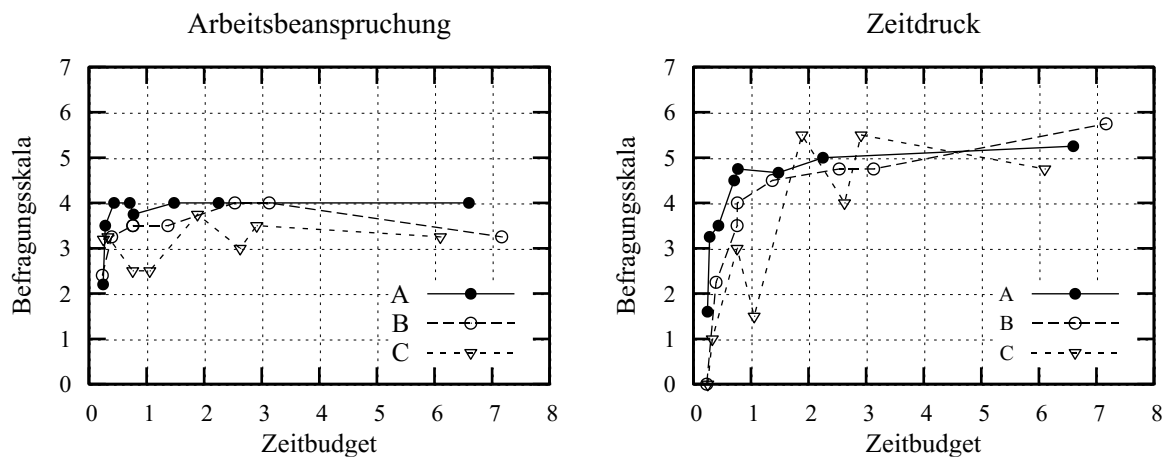


Abbildung 47: Mittlere Arbeitsbeanspruchung pro Intervall über das mittlere Zeitbudget (links). Mittlerer Zeitdruck pro Intervall über das mittlere Zeitbudget (rechts)

Ebenso wie bei dem Zusammenhang zwischen der Häufigkeit an Zeitfehlern und dem Zeitbudget ist der Bereich $\widehat{TB} > 1$ folglich als ein Überlastbereich zu erachten. Insbesondere der Verlauf der Probanden A und B lässt die Vermutung zu, dass der Proband in diesem Bereich nicht befähigt ist, nach weiteren Abstufungen in der gesetzten Last differenzieren zu können. Aufgrund des übereinstimmenden Trends aller drei Probanden, zeichnet sich der empfundene Zeitdruck im Intervall $TB \leq 1$ als steuerbare Folgegröße der Latenzzeit aus. Ebenso wie bei der Reaktion der Zeitfehler im vorhergehenden Abschnitt kann der obere Lastbereich im Intervall $t_L = 20 - 40$ s identifiziert werden.

3.6.4 Die Analyse der Runway Incursion

Die *RI-Rate* als Reaktion der gesetzten Latenzzeit ist anhand einer linearen Regressionsanalyse in Anhang H.7 analysiert worden. Sowohl die Signifikanzwerte als auch die Bestimmtheitsmaße lassen keinen Rückschluss auf einen Zusammenhang zu. Es ist zu vermuten, dass die Qualitätsminderung in Folge des Zeitdrucks innerhalb der Zeitintervalle zu

kurz anhielt, um eine ausreichende Anzahl an Runway Incursion bewirken zu können. Die zusätzliche Kalibrierung anhand der beobachteten Häufigkeit an Runway Incursion kann somit aufgrund mangelnder Stichproben nicht vorgenommen werden.

3.6.5 Die Bestimmung der Zeitdrucklastparameter

Im Folgenden werden die Kriterien für die Bestimmung der Zeitdrucklastparameter der Laststufen ComPerf 1 und 2 diskutiert:

1. *Zielbereich*: Die Laststufenparameter sollten innerhalb des zeitdrucksensitiven oberen Lastenbereichs liegen, um die Primärvarianz zu maximieren. Jedoch sind hierbei auch Lerneffekte bei den Probanden mit in die Festlegung einzubeziehen. Ein möglicher Trainingseffekt während der Versuchslaufzeit könnte zu einer zunehmenden routinierten und höheren Arbeitsgeschwindigkeit führen. Eine nachträgliche Verschiebung des oberen Lastenbereiches könnte die Folge sein, da der Proband zu höheren Zeitdrucklasten ein weniger zeitdrucksensitives Verhalten erlangen könnte. Zu niedrig angesetzte Laststufenparameter würden demgegenüber, wie durch die Befragung festgestellt, niedrige Effektstärken zur Folge haben.
2. *ComPerf 1 und 2*: Der Abstand zwischen den Zeitdrucklastparametern sollte derart groß bemessen sein, dass der Effekt der Änderung des Zeitdrucks auf die Stressreaktionen messbar ist. Die Messung bezieht sich dabei vornehmlich auf die Risikometrik *RI-Rate* und die Anzahl der Zeitfehler *ZF*. Ein zu eng gesetzter Abstand würde zu nicht messbaren Unterschieden der Grundgesamtheit der beiden Größen führen. Ein zu weiter Abstand könnte dazu führen, dass ein Lastenparameter außerhalb des oberen Lastenbereiches liegen könnte.

Tabelle 21: Gewählte Parameter für die Laststufen 1 und 2

<i>Laststufe</i>	<i>Latenzzeit des Fluglotsenagenten [s]</i>
t_{L1}	30
t_{L2}	20

Unter Abwägung der soeben genannten Kriterien mit den jeweiligen Vor- und Nachteilen und unter Einbezug der Analyseergebnisse des Zeitfehlers und des subjektiven Zeitdrucks der vorhergehenden Abschnitte wurden die beiden Laststufenparameter gemäß Tabelle 21 gewählt. Sie orientieren sich an dem oberen Lastbereich, dessen Intervall $t_L = 20 - 40$ s mit Hilfe der Ergebnisse der Kalibrierung geschätzt werden konnte. Durch die Setzung der beiden Größen an die obere Lastengrenze könnte eine Leistungssteigerung aufgrund des weiteren Trainings der Probanden weniger stark ins Gewicht fallen. Der Abstand von 10 s lässt aufgrund der Reaktion der Anzahl an Zeitfehlern einen hinreichenden Effekt für die Primärvarianz durch die Lastenänderung erwarten.

3.7 Die Analyse der Lastenabhängigkeit der Aktivzeiten

Die Baseline und die Zeitdrucklastszenarien bieten für den Probanden gleiche Arbeitsbedingungen mit Ausnahme des durch das Zeitdruckinduktionsverfahren vorgegebenen

Wettbewerbs. Der Vergleich der Leistungsmetrik der Szenarien anhand der Messgröße t_{aktiv} soll Unterschiede und Tendenzen der Zeitdrucklastszenarien aufzeigen. Der Vergleich liefert den Rückschluss, ob sich die Leistung der Flugplatzkontrolle $P_{P\&S}$ signifikant geändert hat und gibt Rückschlüsse auf Lerneffekte, die sich über die Runden aufgebaut haben.

Gemäß des Leistungsmodells wird eine Verkleinerung des Medians der Aktivzeiten \tilde{t}_{aktiv} bei zunehmender Last erwartet: $\tilde{t}_{aktiv\ Last} < \tilde{t}_{aktiv\ BL}$. Es wird ebenso eine Abhängigkeit zu dem Mittelwert vermutet, der durch $\hat{t}_{aktiv\ Last} < \hat{t}_{aktiv\ BL}$ ausgedrückt werden kann. Dieser ist die ausschlaggebende Größe im Leistungsvergleich des Wettbewerbs mit dem Fluglotsenagenten.

Die Abbildung 48 visualisiert hierzu die Boxplots und Whisker je Versuchsdurchführung, unterteilt in Laststufen und die Baseline (hier mit $t_L \rightarrow \infty$ gekennzeichnet), mit wenigstens 160 Stichproben (Verkehrsbewegungen) pro Durchführung. Die Mediane aller Probanden liegen zwischen 400 und 500 s und untere Whisker zwischen 170 und 200 s. Die großen Unterschiede in dem oberen Whisker deuten auf eine große Streubreite der oberen Ausreißer hin und lassen eine Weibull-ähnliche Verteilung nach Vorbild der Verteilung in Abschnitt 3.3.1 vermuten.

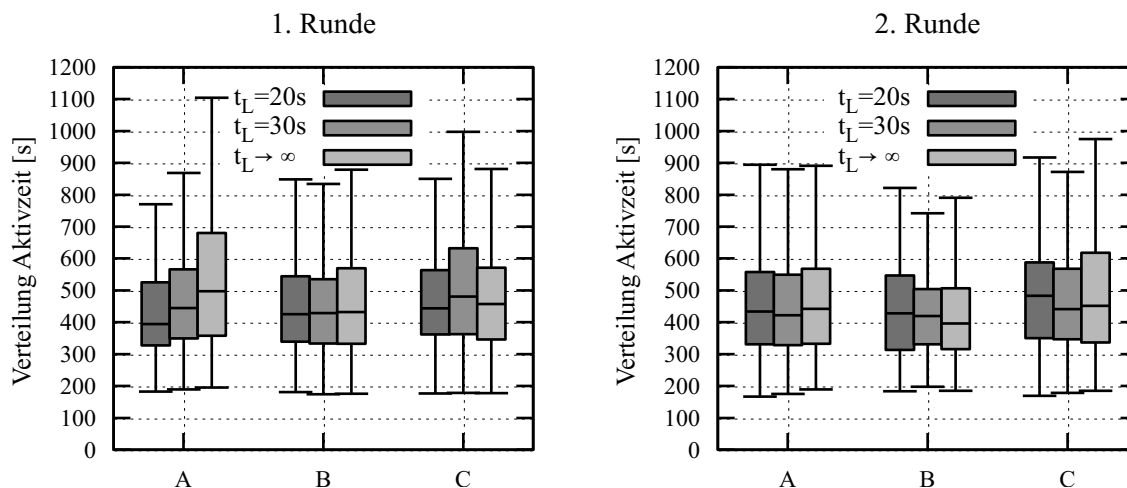


Abbildung 48: Boxplot der Verteilung der Aktivzeiten t_{aktiv}

Die Ähnlichkeit der Aktivzeiten der Baseline zu den Laststufen 1 und 2 deutet darauf hin, dass die Qualität der effizienten Führung invariant zu einer gesetzten Last ist. Der Vergleich der Aktivzeitverteilungen zeigt anhand des statistischen einseitigen Prüfverfahrens U-Test nach Whitney-Mann keine signifikante Verkleinerung der Aktivzeiten bei Lasten gegenüber der Baseline. Dies ist ebenfalls der Fall für die Mittelwertverschiebung anhand eines t-Tests. Die Berechnungen hierzu sind in Anhang I.1 hinterlegt. Der Proband A bildet zur ersten Runde eine Ausnahme. Er weist anhand der Medianverschiebung als auch der Mittelwertverschiebung eine signifikante*** Verkleinerung und somit erwartungskonforme Reaktion der Aktivzeiten auf.

Der geringe Effekt der Latenzzeit auf die Verteilung der Aktivzeiten ist anhand von drei möglichen Ursachen erklärbar:

- *Geringe Effektstärke*: Die Verteilungen der Aktivzeit weisen eine zu große Varianz auf, um die Verschiebung sowohl des Medians als auch des Mittelwertes gegenüber dem Zufall abzugrenzen.
- *Störung durch Zeitfehler*: Eine kleinere Leistungsfähigkeit des Probanden und auch des Fluglotsenagenten führt schließlich zu einem Zeitfehler. Die verbundene Aktivzeit fällt aufgrund des Leistungsdefizits größer aus als erwartet. Je mehr Zeitfehler auftreten, desto größer ist die Aktivzeit. Bei vielen Zeitfehlern kann dieser Effekt signifikant ausfallen. Mit höherer Zeitdrucklast könnte dieser Effekt zunehmend stärker ausfallen.
- *Varianz durch Runway Incursion*: Die Verursachung einer Runway Incursion führt zu einem Wettbewerbsvorteil, da die Aktivzeit durch einen Regelbruch möglicherweise kleiner ausfällt.

Die gemessenen Ergebnisse sind sicherlich das Resultat des Zusammenwirkens aller störenden Ursachen.

3.8 Die Analyse der Lern- und Müdigkeitseffekte

3.8.1 Die Lerneffekte

Bei zunehmendem Training während des Durchführungszeitraumes ist mit einer höheren Leistungsfähigkeit der Probanden zu rechnen. Der Trainingseffekt könnte vermutlich anhand eines Vergleiches der Aktivzeiten beider Runden nachvollzogen werden, wobei in der zweiten Runde kleinere Aktivzeiten zu erwarten sind.

Zum Nachweis dieses Effektes vergleicht ein einseitiger Whitney-Mann-Test die Stichprobensätze zu je gleichen Randbedingungen auf eine mögliche Verschiebung des Medians der Verteilungen gemäß $H_0: \tilde{t}_{aktiv R2} \geq \tilde{t}_{aktiv R1}$. Das Ergebnis zeigt anhand der neun Vergleiche in Anhang I.9 bei

- der Baseline des Probanden A und
- der ersten Laststufe, Proband C,

einfache bis sehr signifikante kleinere Aktivzeiten zur zweiten Runde. Dies kann anhand der Boxplots-Abbildung 48 nachvollzogen werden. Die übrigen Verteilungen deuten auf gleiche Tendenzen hin. Es überwiegt somit die Gleichheit der Aktivzeiten. Dies deutet auf eine mehrheitlich stabilisierte Arbeitsweise hin, bei der die Probanden eine weitestgehend abgeschlossene Lernkurve über die effiziente Kontrolle des Flugverkehrs aufweisen.

3.8.2 Die Erschöpfungseffekte

Bei einer Versuchszeit von ca. 1,5 Stunden pro Versuchsdurchführung und einer erhöhten Arbeitsbelastung ist bei der Arbeit der Novizenprobanden mit einsetzender Müdigkeit und Erschöpfung zu rechnen. Ein daraus resultierender Effekt könnte sich möglicherweise auf die aufgezeichneten Stressreaktionen auswirken. Dies könnte zu einer über die Zeitdauer des Versuches nachlassenden Aufmerksamkeit und Konzentration für zu treffende Entscheidungen führen und eine Minderung der Qualität der Handlungen bewirken. Eine zeitabhängige Varianz der menschlichen Leistungsfähigkeit ist zu prüfen.

Für einen Test bieten sich prinzipiell alle Größen an, die in Verbindung mit der Leistungsfähigkeit des Operators stehen. Für die Überprüfung des Zeitbezugs einer Größe

eignet sich die einfache lineare Regression, die eine Zeitvarianz anhand des P-Wertes und des Bestimmtheitsmaßes belegen kann. Für eine Indikation der Erschöpfung bietet sich die Reaktionszeit an, die bei menschlichen Tätigkeiten eine einfache Abschätzung über die Existenz eines Erschöpfungseffektes liefern kann. Die Ergebnisse des Tests der Reaktionszeiten aller Zeitdrucklastszenarien und Probanden sind in Anhang I.10 hinterlegt. Sie belegen anhand ihrer P-Werte, dass kein signifikanter Effekt festzustellen ist. Die Bestimmtheitsmaße liefern korrespondierende Werte bis maximal 0,5 %. Es kann zwischen der Zeitdauer der Simulation und der Reaktionszeit kein Zusammenhang festgestellt werden. Die Effekte einsetzender Müdigkeit und Erschöpfung werden aus diesem Grund nicht in die weiteren Untersuchungen einbezogen.

3.9 Die Analyse der Stressreaktionen

Die Auswahl der Größen, die hinsichtlich der Stressreaktion analysiert wurden, wurde den Kausalhypothesen in Abschnitt 3.1.2 angepasst. Die Messergebnisse aller Versuchsdurchführungen sind in Anhang I.3 tabellarisch zusammengefasst. Zusätzlich wurde eine Korrelationstabelle in Anhang I.4 erstellt, die die Korrelationskoeffizienten nach der Methode von Pearson auf Basis der Messgrößen des Anhangs I.3 enthält. Die Korrelationstabelle wurde um eine Spalte für das angepasste Zeitbudget ergänzt, deren Werte anhand der Definition t_r/t_L (Formel 3.17) berechnet wurden. Ebenso wurde eine Spalte für die relative Häufigkeit der Runway Incursion pro Startfreigabe (*RI-Rate*) hinzugefügt. Anhand der Farbcodierung gibt die Tabelle eine einfache Übersicht über mögliche Zusammenhänge mit hohem positivem und mit negativem Korrelationsfaktor. Zusätzlich werden für die Analyse der unterstellten Zusammenhänge statistische Tests verwendet, die im Folgenden vorgestellt werden.

3.9.1 Die Reaktionszeit

Die Verteilungen der Reaktionszeit t_r der jeweiligen Lastenstufe und Runde wurde mit Hilfe eines Mann-Whitney-U-Tests auf einen möglichen Effekt der Zeitdrucklast getestet. Gemäß dem angepassten Risikomodell ist eine Abnahme der Reaktionszeit bei zunehmender Zeitdrucklast zu erwarten. Entsprechend der Kausalhypothese wird die Alternativhypothese formuliert, die die Verkleinerung bei zunehmender Last formuliert: $H_1: \tilde{t}_{r\ CP2} < \tilde{t}_{r\ CP1}$. Die Tabelle 22 zeigt die Ergebnisse des einseitigen U-Tests nach Runde und Probanden aufgeschlüsselt. Die entsprechenden Rechnungen sind dem Anhang I.7 zu entnehmen. Beim Vergleich fällt auf, dass Proband B in der zweiten Runde sehr signifikant auf die zunehmende Last mit einer abnehmenden Reaktionszeit reagiert hat. Dem Probanden A in der zweiten Runde und Proband C in der ersten Runde sind ebensolche Zusammenhänge zu attestieren, die aber nicht das Kriterium der Signifikanz erfüllen, jedoch mit dem P-Wert unterhalb von 10 % liegen.

3.9.2 Die Aktivzeit

Die Aktivzeiten beider Runden und aller Probanden sind in Abbildung 48 (S. 120) gegenübergestellt. Gemäß den Kausalhypothesen dieser Pilotstudie ist eine Verkleinerung der Aktivzeiten bei zunehmender Last zu erwarten. Dieses lässt sich anhand eines einseitigen U-Tests nach Whitney-Mann nachprüfen, mit der Alternativhypothese $H_1: \tilde{t}_{aktiv\ CP2} < \tilde{t}_{aktiv\ CP1}$.

Die Ergebnisse des Tests (Anhang I.7) liefern keine signifikante Verkleinerung des Medians bei zunehmender Last, mit Ausnahme der Probanden A und C innerhalb der ersten Runde. Wie in Abschnitt 3.7 bereits vermutet ist der eindeutige Trend der Verschiebung des Medians aufgrund der geringen Effektstärke nicht bestimmbar. Zusätzliche Störeffekte durch die Zeitfehler und die Verursachung von Runway Incursion können ebenfalls eine nicht nachvollziehbare Verschiebung bewirken.

Tabelle 22: P-Werte des einseitigen Mann-Whitney-U-Tests auf Gleichheit oder Vergrößerung der Reaktionszeit bei zunehmender Last

Proband		P-Wert [%]		
		A	B	C
Runde	1	51,1	43,7	9,3
	2	6,1	0,9**	91,6

3.9.3 Das angepasste Zeitbudget

Der Zusammenhang zwischen Latenzzeit und dem resultierenden angepassten Zeitbudget mit der Definition t_r/t_L (Formel 3.17) ist signifikant über alle Probanden, Runden und Laststufen gemäß des Signifikanztests, bei dem die lineare Regression ein Bestimmtheitsmaß von 66,42 % (P-Wert: 0,12 %) ** liefert (Anhang I.12). Es kann somit Linearität zwischen der angelegten Last und dem resultierenden Zeitbudget angenommen werden.

3.9.4 Die Arbeitsbeanspruchung und der subjektive Zeitdruck

Die Befragungsergebnisse der zwei Laststufen ComPerf 1 und 2 wurden nach Runde 1 und 2 sowie Probanden aufgeschlüsselt und in Abbildung 49 als Mittelwerte mit Standardabweichung als oberer und unterer Whisker gegenübergestellt. Eine derartige Darstellung unterstellt eine Standardverteilung der Befragungsergebnisse. Der Einfluss einer zunehmenden Zeitdrucklast bewirkt vermutlich eine Zunahme des empfundenen Zeitdrucks auf der Befragungsskala. Somit lautet die entsprechende Alternativhypothese $H_1: \tilde{x}_{CP1} < \tilde{x}_{CP2}$. Das Ergebnis des einseitigen Mann-Whitney-Tests, aufgeführt in Anhang I.6, liefert keine signifikanten P-Werte. Es ist somit von keiner Abhängigkeit des empfundenen Zeitdrucks zu der gesetzten Laststufe auszugehen. Bei Unterstellung einer Standardverteilung bestätigt dies das Ergebnis eines T-Test.

Die Arbeitsbeanspruchung hingegen sollte erwartungsgemäß keinen Effekt durch die gesetzte Zeitdrucklast erfahren. Die entsprechende Nullhypothese lautet $H_0: \tilde{x}_{CP1} = \tilde{x}_{CP2}$. Der beidseitige U-Tests nach Mann-Whitney (Anhang I.5) liefert keinen signifikanten P-Wert, der auf eine Ablehnung der These gleicher zentralen Tendenzen schließen lässt. Bei Unterstellung einer Standardverteilung der Befragung kommt der zweiseitige T-Test zu einem ebenso nicht-signifikanten Ergebnis für alle Stichprobensätze.

Die Ergebnisse der Befragung, sowohl Arbeitsbeanspruchung als auch Zeitdruck, lassen keine eindeutige Gesamttendenz erkennen. Dies steht in Widerspruch zu den Ergebnissen der Kalibrierung 3.6.3, die einen signifikant-linearen Zusammenhang aufweist. Bei Betrachtung der Zeitdrucklast beläuft sich der Unterschied bei der Kalibrierung auf eine unmittelbare Abfolge von verschiedenen Laststufen, die systematisch durchfahren werden. Im

Gegensatz dazu variiert die Last der Hauptversuche während der Versuchsdurchführung nicht. Nach Tagen oder gar Wochen zwischen den Versuchsdurchführungen fehlt dem Probanden das Gefühl, unterschiedliche Zeitdrucklasten unterscheiden zu können. Die Folge ist eine angepasste Wahrnehmung der Zeitdrucklasten.

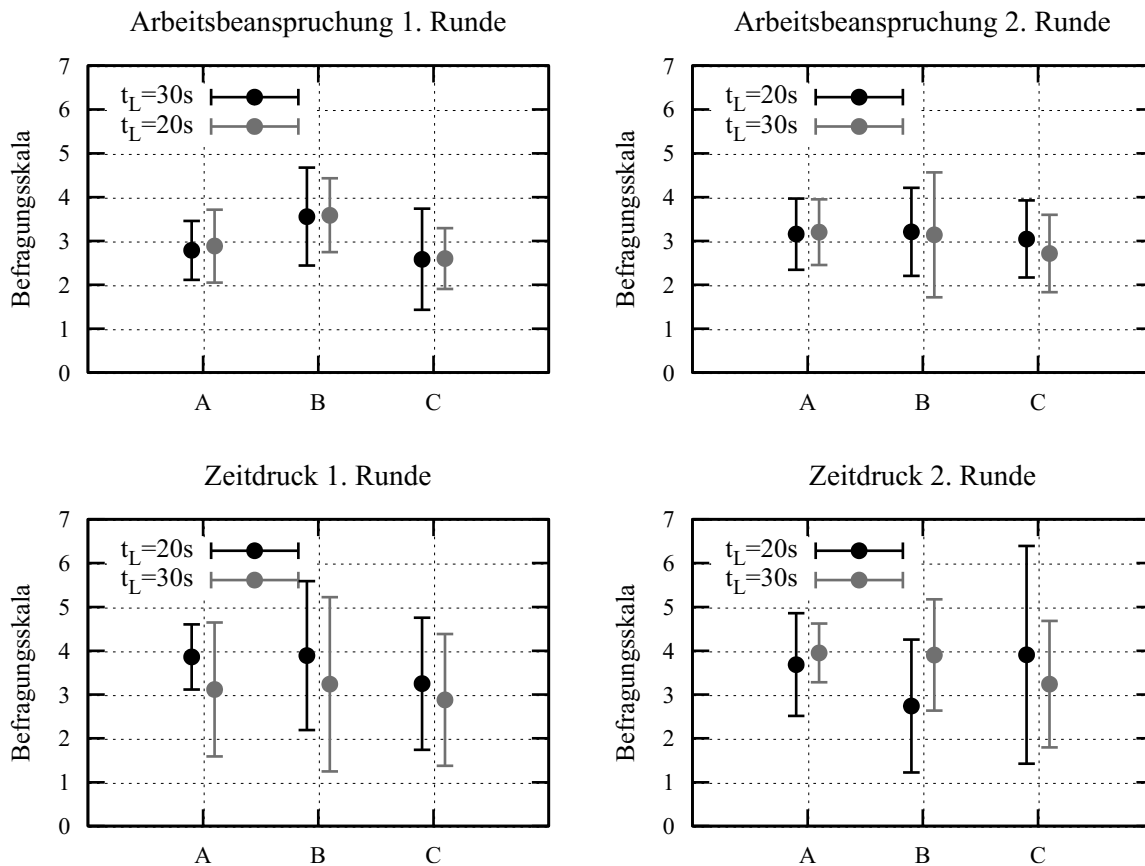


Abbildung 49: Arbeitsbeanspruchung und Zeitdruck-Befragung beider Runden in der Übersicht

3.9.5 Die Runway Incursion und der Zeitfehler

Die *RI-Rate* und Häufigkeit an Zeitfehlern (*ZF*) der zwei Laststufen ComPerf 1 und 2 wurden nach Runde 1 und 2 sowie Probanden aufgeschlüsselt und gegenübergestellt. Die Übersicht in Abbildung 50 zeigt die Gegenüberstellung von *RI-Rate* und auch die Häufigkeit von Zeitfehlern bei Variation der Zeitdrucklast anhand einer Vektordarstellung im XY-Diagramm.

Reaktion der *RI-Rate* und Anzahl der Zeitfehler nach Runden

Die lineare Regressionsanalyse (Anhang I.14) liefert zwischen der Latenzzeit und der *RI-Rate* ein Bestimmtheitsmaß von 11,1 % (P-Wert 24,4 %) einen nicht signifikanten Zusammenhang. Ebenso liefert der Test für die Latenzzeit zur Häufigkeit von Zeitfehlern mit einem Bestimmtheitsmaß von 8,3 % (P-Wert 31,5 %) einen nicht signifikanten Zusammenhang. Wird nach den zwei Runden unterschieden, so resultiert ebenfalls kein signifikanter Zusammenhang (Tabelle 23).

Nach Auswertung der Daten kann nicht von einer Intensivierung der *RI-Rate* und der Zeitfehler in Abhängigkeit der Latenzzeit mit den in der Tabelle ausgewiesenen P-Werten ausgegangen werden. Bei Betrachtung der Vektoren in Abbildung 50 fällt eine relativ starke

Verteilung der Arbeitspunkte der Probanden auf. Jeder Proband zeigt eine Reaktion auf die gesteigerte Last, die bis auf den Probanden B in der ersten Runde in der Länge ähnlich ausfällt, sich aber in der Richtung unterscheidet.

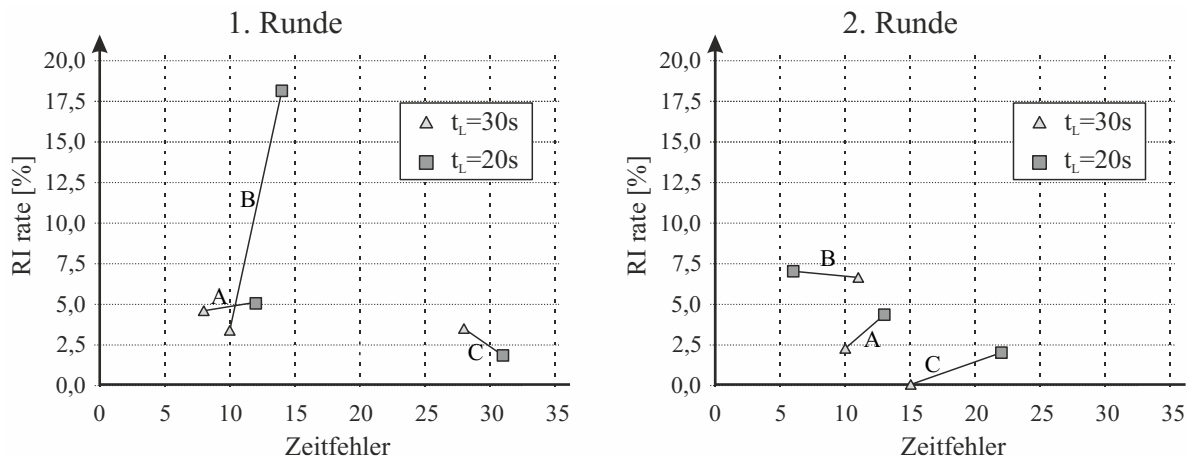


Abbildung 50: Runway Incursion-Rate und Zeitfehlerhäufigkeiten als lastenabhängige Betriebspunkte der Latenzzeit

Die Reaktion der RI-Rate und Anzahl der Zeitfehler bei Zusammenführung der Runden

Die Zusammenführung der Runden und der Stichproben aller Probanden lässt eine höhere Stichprobenzahl und somit eine größere statistische Teststärke im Ergebnis erwarten. Für die Zusammenführung wurde die *RI-Rate* beider Runden entsprechend ihrer relativen Werte gemittelt. Die Anzahl der Zeitfehler wurde addiert. Das Ergebnis der Zusammenführung ist in Abbildung 51 dargestellt, die die Stressreaktionen der drei Testpersonen als Vektoren in der Übersicht zeigt.

Tabelle 23: Bestimmtheitsmaße der Zusammenhänge Latenzzeit t_L , *RI-Rate* und Zeitfehler *ZF*

Runde	Bestimmtheitsmaß [%] (P-Wert [%])	
	t_L - <i>RI-Rate</i>	t_L -Zeitfehler
1	17,2 (41,4)	4,1 (69,7)
2	10,6 (43,1)	13,3 (37,4)

Die Abbildung zeigt die lastenabhängige Stressreaktion in vektorieller Darstellung pro Proband. Deutlich zu erkennen sind die Stressreaktionen bei Proband A und C, die sowohl in der *RI-Rate* als auch bei der Häufigkeit von Zeitfehlern eine Zunahme bei steigender Last erfahren. Dies ist ein plausibles Ergebnis des Effektes zunehmenden Zeitdrucks. Es stimmt mit den Annahmen über die Wechselwirkung zwischen Risikobereitschaft und Schnelligkeit, eingeführt in Abschnitt 3.2.5.2, überein.

Ein nachteiliger Effekt der Zusammenführung beider Runden ist die Dominanz des Ergebnisses durch Ausreißer. Dieser Effekt kann bei Kandidat B beobachtet werden, dessen bessere Sicherheitsleistung in der Runde 2 durch den Ausreißer in der ersten Runde überzeichnet wird. Proband B erfährt in beiden Runden eine zunehmende *RI-Rate*, die zur Erwartung des Konzeptes AccSis steht. Die Abnahme der Häufigkeit von Zeitfehlern ist

hingegen nicht erwartungskonform, da dieser Proband eine gegenläufige Tendenz bei steigender Last aufweist. Diese Reaktion bestätigt allerdings die Analyse der Reaktionszeit t_r in Abschnitt 3.9.1, die eine sehr signifikante Abnahme der Reaktionszeit ergab. Der daraus gewonnene Vorsprung bewirkt eine Verringerung der Zeitfehler. Allerdings ist das Ausmaß der Reaktion in der *RI-Rate* gegenüber den anderen Probanden relativ stark ausgeprägt. Insbesondere in Runde 1 könnte es sich hier um einen Fall von Überlastung gehandelt haben. Dabei könnte der Proband vorsätzlich oder unbewusst Runway Incursion erzeugt haben, um Zeitfehler zu vermeiden. Es entstünde somit ein Vorteil im Wettbewerb gegenüber dem Fluglotsenagenten. Vielleicht ist dies auch der Reihenfolge des Versuchsplans geschuldet, die mit der hohen Last begann. Möglicherweise unterlag der Proband Eingewöhnungsproblemen, im Wettbewerb zu arbeiten. Dies erscheint möglich, da das Arbeiten unter Einwirken des Zeitdruckinduktionsverfahrens nicht Bestandteil des Trainings und der Qualifikation war.

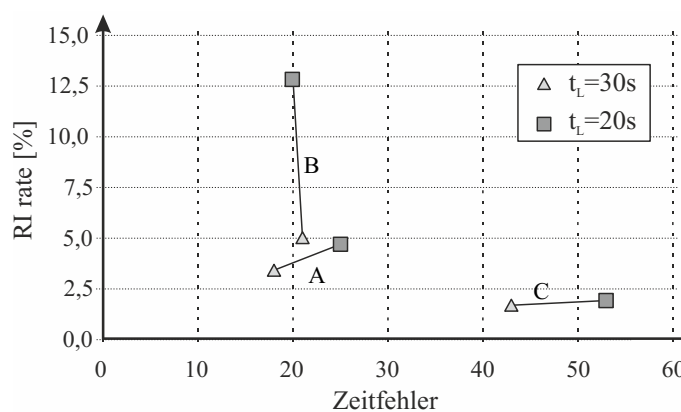


Abbildung 51: Zusammenführung der *RI-Rate* und Zeitfehler-Häufigkeiten beider Runden als lastenabhängige Betriebspunkte der Latenzzeit

Auffällig ist das Rechteck des Zeitfehlers und *RI-Rate* von Proband A und C, deren Seitenverhältnisse sich bei dem jeweiligen Probanden über die Laststufen anscheinend konstant halten. Eine Systematik könnte hier unterstellt werden. Die Wechselwirkung zwischen *RI-Rate* und Zeitfehler wird anhand einer Korrelationsanalyse geprüft. Das Ergebnis ist ein Korrelationskoeffizient von -57,7 % bei einem P-Wert von 22,9 % und liefert den Anhaltspunkt für einen Zusammenhang. Die Berechnung ist in Anhang I.13 hinterlegt. So scheint insbesondere das Verhalten von Proband B nicht zu einem linearen Zusammenhang zwischen den beiden Größen zu passen. Aber auch Proband A und C variieren deutlich in dem Versatz der beiden Messgrößen. Dies lässt die Vermutung zu, dass neben der Wechselwirkung zwischen der Risikobereitschaft und Schnelligkeit weitere Effekte wirken, die aus dem individuellen Arbeitsverhalten des Probanden entstehen.

Analyse mit Hilfe des Konzeptes der Gesamtleistung

Wenngleich auch die Stressreaktionen zunächst unkontrolliert wirken, lassen die in Abbildung 50 und auch in Abbildung 51 gezeigten Stressreaktionen systematische Muster der Reaktion

erahnen. Für die Ergründung eines dazu passenden Erklärungsansatzes könnten nach Analyse die zwei folgenden Faktoren einen entscheidenden Einfluss darstellen:

- *Proportionierung*: Abwägung zwischen Risikobereitschaft und Schnelligkeit, beschrieben durch das erweiterte Risikomodell.
- *Individuelle Basisleistung*: Der Versatz durch variierende Basisleistungen des Probanden zur Basislast, erläutert in Abschnitt 2.5.4.3.

Eine Berücksichtigung der damit verbundenen Effekte könnte die Systematik der beobachteten Stressreaktion eingrenzen. Die Adaption der beiden Faktoren auf den vorliegenden Kontext erfolgt in den folgenden Arbeitsschritten:

1. *Konzept der Gesamtleistung*: Die Kombination des Risikos und der Schnelligkeit zu einer Gesamtleistung stellt einen Ansatz zur Berücksichtigung des Abwägungsprozesses dar. Aufgrund der gleichmäßigen Proportionen zwischen Zeitfehlern und *RI-Rate*, beobachtet in Abbildung 51, wird die Multiplikation beider Leistungsmetriken gemäß $P_{gesamt} = P_{konfliktfrei} \cdot P_{P\&S}$ gewählt. Die Multiplikation erscheint konzeptionell plausibel, da eine Gesamtleistung nur vorhanden sein kann, wenn sowohl die Sicherheitsleistung als auch die Schnelligkeit beide ungleich Null sind.
2. *Substitution der benötigten Zeit*: Erachtet man die benötigte Zeit t_n als die Reaktionszeit t_r , so kann die Definition des angepassten Zeitbudgets gemäß Formel 3.17 eingesetzt werden.
3. *Substitution der Zeitfehler*: Der Korrelationskoeffizient nach Pearson von 56,7 % (P-Wert 3,44 %) verdeutlicht einen signifikant* linearen Zusammenhang zwischen dem angepassten Zeitbudget (vgl. Formel 3.17) und der Häufigkeit an Zeitfehlern *ZF* (siehe Anhang I.4, Korrelationsmatrix nach Pearson). Ein Substitution von \widehat{TB} durch *ZF* erscheint somit für weitere Zusammenhangsbetrachtungen legitim.
4. *Konzept der Basisleistung*: Die Berücksichtigung der Basisleistung P_0 erfolgt anhand einer Normierung mittels Division der Basisleistung am Arbeitspunkt $t_L = 30s$.

Die Leistung $P_{P\&S}$ kann mit Hilfe der Formel 3.7 in eine zeitliche und eine qualitative Komponente zerlegt werden. Bei Einsetzen der Sicherheitsleistung $P_{konfliktfrei}$ wiederum kann nach den zwei möglichen Ansätzen der Definition unterschieden werden. Diese sind in Abschnitt 3.2.5.2 in den Formeln 3.5 und 3.6 vorgestellt. Die beiden Ansätze werden im Folgenden vorgestellt und deren Gesamtleistung entsprechend den soeben vorgestellten Arbeitsschritten in jeweils einen Term überführt. Diese sollen je einen Zusammenhang beschreiben, der anhand der linearen Regressionsanalyse auf Bestimmtheitsmaß und Signifikanz getestet wird.

- *Das Risiko als Komplementärwahrscheinlichkeit der Sicherheitsleistung*: Durch Einsetzen der *RI-Rate* als Komplementärwahrscheinlichkeit und Umsetzung der oben genannten Punkte entsteht:

$$\begin{aligned}
 P_{gesamt} &= P_{konfliktfrei} \cdot P_{P\&S} = (1 - RI_{rate}) \cdot \frac{Q_{P\&S}}{t_r} \\
 &= (1 - RI_{rate}) \cdot \frac{Q_{P\&S}}{\overline{TB} \cdot t_L} \\
 &= Q_{P\&S} \cdot \frac{(1 - RI_{rate})}{ZF \cdot t_L} \\
 \Leftrightarrow t_L &= \frac{Q_{P\&S}}{P_{gesamt}} \cdot \frac{(1 - RI_{rate})}{ZF}
 \end{aligned} \tag{3.18}$$

Der Zusammenhang zwischen der Latenzzeit und der beobachteten *RI-Rate* und Zeitfehler resultiert somit zu

$$t_L \sim \frac{1 - RI_{rate}}{ZF}, \tag{3.19}$$

Die nun folgende Normierung liefert auf Basis des Arbeitspunktes bei $t_L = 30s$

$$t_L \sim \frac{1 - RI_{rate}}{ZF} \cdot \frac{ZF(t_L = 30s)}{1 - RI_{rate}(t_L = 30s)}. \tag{3.20}$$

Die lineare Regressionsanalyse (Anhang I.15) liefert mit einem Bestimmtheitsmaß von 40,5 % bei einem P-Wert von 17,4 % keinen signifikanten Zusammenhang.

- *Das Risiko als Kehrwert der Sicherheitsleistung:* Durch Einsetzen der *RI-Rate* als Kehrwert und Umsetzung der genannten Arbeitsschritte entsteht

$$P_{konfliktfrei} \cdot P_{P\&S} = Q_{P\&S} \cdot \frac{1}{RI_{rate} \cdot ZF \cdot t_L}. \tag{3.21}$$

Im Folgenden entsteht ein Zusammenhang zwischen der Latenzzeit und den beobachteten Größen zu

$$t_L \sim \frac{1}{RI_{rate} \cdot ZF}. \tag{3.22}$$

Es entsteht nach Normierung durch den Arbeitspunkt bei $t_L = 30s$ als Basisleistung

$$t_L \sim \frac{RI_{rate}(t_L = 30s) \cdot ZF(t_L = 30s)}{RI_{rate} \cdot ZF}. \tag{3.23}$$

Die lineare Regressionsanalyse (Anhang I.15) liefert mit einem Bestimmtheitsmaß von 70,1 % und einem P-Wert von 3,75 % einen signifikanten* Zusammenhang.

Die Zusammenfassung der gemessenen Metriken als Gesamtleistung liefert eine mögliche Erklärung der Varianz der individuellen Stressreaktionen. Die Reaktion des Probanden kann hierbei innerhalb eines Ermessensspielraums liegen. Erst bei einer Zusammenfassung der individuellen Freiheitsgrade kann eine gemeinsame Systematik erkannt werden, mit der der Proband auf die Zeitdrucklasten reagiert.

Bestimmung der Effektstärke nach Cohen

Unter der Annahme einer Verteilung ähnlich der der Standardverteilung ist die Effektstärke nach Cohen gemäß

$$d = \frac{X_{n20} - X_{n30}}{\sigma_{Pooled}} \quad (3.24)$$

ermittelbar. So kann d mit Hilfe der Anzahl der Runway Incursion pro Versuchsdurchführung auf 0,67 bestimmt werden. Dies entspricht nach gängiger Klassifizierung nach Cohen (Bortz & Döring, 2006) einem mittleren (0,5) bis großen Effekt (über 0,8). Die dazugehörige Rechnung zur Ermittlung des Wertes d ist in Anhang I.16 hinterlegt.

3.10 Diskussion der Ergebnisse

Dieser Abschnitt hat das Ziel, die zu Beginn dieses Kapitels erläuterten Fragestellungen zu beantworten und Aussagen über die Effektivität des Zeitdruckinduktionsverfahrens und damit verbundenen Bedingungen zu schlussfolgern. Die Ergebnisse der drei Probanden zeigen Reaktionen der menschlichen Arbeitsweise, die wahrscheinlich in einem Zusammenhang mit dem induzierten Zeitdruck stehen. Die dabei unterstellten Kausalzusammenhänge konnten teilweise anhand von Signifikanztests am Probanden und probandenübergreifend nachvollzogen werden. Im Folgenden werden die initialen Fragestellungen dieser experimentellen Pilotstudie den Analyseergebnissen gegenübergestellt.

Sensitivität des Risikos zum induzierten Zeitdruck

Das Hauptergebnis dieses Kapitels stellt die Empfindlichkeit der *RI-Rate* und Zeitfehler gegenüber der gesetzten Last dar. Der Effekt der gesetzten Latenzzeit auf eine Intensivierung der *RI-Rate* ist nicht bei allen Probanden eindeutig festzustellen (Anhang I.14).

Die Zusammenfassung der Risikometrik und der Zeitfehler zu einer Gesamtleistung des Probanden und die anschließende Normierung bieten demgegenüber einen vielversprechenderen Erklärungsansatz zum Verstehen und zur Beschreibung der zu erwartenden Stressreaktion. Die lineare Regressionsanalyse ergab für zwei unterschiedliche Modellierungsansätze der Gesamtleistung

- ein Bestimmtheitsmaß von 40,5 % bei einem P-Wert von 17,4 % für das Risiko als Komplementärwahrscheinlichkeit der Sicherheitsleistung und
- ein Bestimmtheitsmaß von 70,1 % bei einem P-Wert von 3,75 % für das Risiko als Kehrwert der Sicherheitsleistung.

Somit liefert der Kehrwertansatz eine Erklärung für einen überwiegenden Teil der Varianz, die den individuellen Stressreaktionen der drei Probanden geschuldet ist. Zusätzlich konnte der Effekt des zunehmenden Zeitdrucks auf die *RI-Rate* anhand des Modells der Binomialverteilung auf einen P-Wert von 1,2% bestimmt werden. Innerhalb der Pilotstudie weist dieses Ergebnis die Gültigkeit der Latenzzeit als effektive Zeitdrucklast für die drei Probanden aus.

Diesem Ergebnis stehen die Ergebnisse der subjektiven Einschätzung der Arbeitsbeanspruchung und des Zeitdrucks durch den Probanden gegenüber, für die in Abschnitt 3.9.4 keine Reaktion in Folge der variierten Zeitdrucklast festgestellt werden konnte. Dies kann möglicherweise anhand der zeitlichen Abstände, mit der die Durchführungen der Zeitdrucklastszenarien geplant und durchgeführt wurden, erklärt werden (vgl. Zeitplan der Durchführungen, Anhang F). Die zeitlichen Abstände zwischen den

Durchführungen führen zu der Schwierigkeit der Probanden, die empfundenen Zeitdrucklasten mit anderen Lasten vergleichen und in Bezug setzen zu können. Aus diesem Grund fallen die Antworten sehr ähnlich aus.

Wechselwirkung mit der Basislast

Die Bestimmung der Grenzwertgrößen der Basislast t_{a0} und t_{n0} ist gemäß Abschnitt 2.6.2 die Voraussetzung für die Bestimmung der Intensität a zu den jeweilig gesetzten Zeitdrucklasten. Die Wechselwirkung zwischen der Basislast und dem induzierten Zeitdruck kann somit für das gezielte Anfahren von Zeitdrucklasten kalibriert werden. Insbesondere bieten sich hier die

- die Reaktionszeit und
- die Aktivzeit

als geeignete Größen an.

Allerdings lieferten weder die Aktivzeiten noch die Reaktionszeiten bei dem Vergleich der Baseline mit den Laststufen der Zeitdrucklastszenarien einen eindeutigen Trend. Die Ursachen hierfür sind unbekannt. Die folgenden Faktoren geben Hinweise auf Störfaktoren, deren Einfluss nicht im genügenden Maße bekannt ist:

- Obwohl die Stichprobenzahlen pro Durchführung bei min. 160 Aktivzeiten bzw. min. 500 Reaktionszeiten liegen, ist die Sekundärvarianz gegenüber der Primärvarianz zu groß.
- Der Effekt von Zeitfehlern und Runway Incursion kann zu einer Verschiebung der Aktivzeiten führen. Dieser Effekt wurde bereits in Abschnitt 3.7 diskutiert.

Maximale Intensivierung der Unfallwahrscheinlichkeit

Die theoretisch maximal mögliche Intensivierung wurde während der Kalibrierung bei Bedingungen zu ($t_L = 0s$) gefahren. Da kein Bezug zum Designstress hergestellt werden konnte, ist die angefahrene Zeitdrucklast nicht kalibriert. Aus diesem Grund ist die maximale Intensivierung nicht ermittelbar. Für die Ermittlung des Designstresses hätten sowohl die Aktivzeit als auch die Reaktionszeit dienen können, die allerdings keinen Zusammenhang zur Zeitdrucklast erkennen lassen. Es ist zu vermuten, dass es sich hier um ein Problem der Repräsentativität der gewählten Größen handelt. Eine weitere Entwicklung von geeigneten Indikatoren zur Erfassung der Stressreaktionen ist notwendig.

Individualvarianz der Ergebnisse

Die ermittelten Stressreaktionen lassen Varianzen erkennen, die vermutlich auf individuelle Personenmerkmale zurückgeführt werden können:

- *Risikobereitschaft und Schnelligkeit*: Die Freiheitsgrade des Probanden, zwischen dem Risiko und der Schnelligkeit abzuwägen, ist nach Analyse der Stressreaktion ein beobachtbarer Effekt. Es wird ein Personenmerkmal deutlich, das sich mit Hilfe des Zeitdruckinduktionsverfahrens nicht kontrollieren lässt.
- *Varianz der Basisleistung*: Die unterschiedliche Basisleistung des jeweiligen Individuums stellt einen systematischen Störeffekt dar, der vermutlich in der unterschiedlichen Leistung zur Basislast unter Designstressbedingungen begründet ist.

Aufgrund der Verwendung von drei Probanden kann kein verallgemeinerter probandenübergreifender Trend festgestellt werden. Die Ergebnisse der Pilotstudie sind aus diesem Grund vorläufige Ergebnisse, die keinen Rückschluss auf die Effektivität des Zeitdruckinduktionsverfahrens bei Austausch der Probanden erlauben. Jedoch zeigen die Vergleiche der Aktivzeiten in Abschnitt 3.7 und Abschnitt 3.9.2 ähnliche Verteilungen für alle Probanden, die auf eine ähnliche Leistungsfähigkeit hinweist. Dies stützt gemäß der Vermutung in Abschnitt 3.4.1 die Annahme einer ähnlichen Leistungsfähigkeit bei Verwendung vorselektierter Versuchsgruppen.

4 Abschlussdiskussion

4.1 Das Fazit über das Konzept „Accelerated Risk Analysis“

Das Ziel, das Risiko und die zugehörigen Risikoschwerpunkte eines neuartigen soziotechnischen Systems mit Hilfe von AccSis zu bestimmen, konnte in seiner Gänze innerhalb dieser Arbeit erwartungsgemäß nicht demonstriert werden. So konnte die experimentelle Studie demonstrieren, den Zeitdruck am Beispiel von 3 Novizenprobanden als Stressfaktor zu verwenden und die Sicherheitsmargen zu schmälern. Die Vermutung, dass AccSis bezüglich der Lösung der Rechtszensur effektiv ist, ist anhand der Ergebnisse unter Berücksichtigung der experimentellen Bedingungen bestätigt, da eine erhöhte Anzahl von Runway Incursion festgestellt werden konnte. Anhand der Systematik der Reaktion kann davon ausgegangen werden, dass die Stressreaktion kontrolliert beeinflusst werden kann und die Intensität aus diesem Grund unter Berücksichtigung der menschlichen Varianzen reproduzierbar ist. Eine Einschränkung für diese Aussage betrifft das Kriterium der Repräsentativität. Mit 3 Probanden ist kein Rückschluss auf die Gesamtheit der Population derart ausgebildeter Probanden gestattet. Allerdings konnten die Ergebnisse Tendenzen aufzeigen, welche Varianzen bei den Stressreaktionen zu erwarten sind. Dies konnte z.B. anhand des Zusammenhangs zwischen Zeitfehlern und Risiko gezeigt werden, was anhand des Modells der Gesamtleistung eine Erklärung für einen Teil der Varianz lieferte. Im Folgenden können zu den Verifikationszielen des Konzeptes AccSis die folgenden Aussagen getroffen werden:

- *Reproduktion des empfundenen Zeitdrucks*: Die Fähigkeit, die Zeitdrucklast über die Zeit konstant zu halten, konnte anhand der systematischen Reaktion von den Probanden A und C bestätigt werden. Allerdings begründet sich ein Vorbehalt bei dieser Aussage in der Stressreaktion, die bei Proband B sowohl in Runde 1 als auch Runde 2 nichterwartungskonforme Richtungen aufzeigte. Obwohl ein einzelner Proband als Ausreißer gelten könnte, zeigt seine Reaktion, welche Bandbreite individuelle Reaktionen haben können. Mögliche Erklärungsmodelle zu dieser Varianz werden im weiteren Verlauf dieses Kapitels vorgestellt. Eine weitere mögliche Varianz bei der Reaktion betrifft die zeitliche Abhängigkeit der Wahrnehmung des Zeitdrucks durch den Probanden. Diese kann anhand der Reaktionen zu den jeweiligen Runden nachvollzogen werden. Eine Minderung der Effektivität des Zeitdrucks bzw. zunehmende Unempfindlichkeit würde erwartungsgemäß eine Verschiebung der Seitenverhältnisse im *RI-Rate-Zeitfehler-Diagramm* zugunsten einer höheren Zeitfehlerhäufigkeit bewirken. Eine derartige Verschiebung erscheint bei Proband A anhand der Stressreaktion möglich. Allerdings gilt es hier zu berücksichtigen, dass die Intensität der Reaktionen von *RI-Rate* als auch Zeitfehler insgesamt bei allen Probanden kleiner ausfallen. Eine wahrscheinlichere Erklärung ist die Gesamtleistung aller Probanden, die zur Runde 2 zugenommen hat. Das Training hat die Probanden befähigt, effektiver und sicherer zu arbeiten. Sowohl die individuelle Stressreaktion als auch der Trainingseffekt mit der dazugehörigen Leistungszunahme überzeichnen an

dieser Stelle die Auswirkung einer möglichen Zunahme der Unempfindlichkeit gegenüber dem induzierten Zeitdruck.

- *Kalibrierung des Zeitdrucks*: Die Grenzwertgrößen konnten weder anhand der Aktivzeit noch anhand der Reaktionszeit bestimmt werden. Mögliche Erklärungen wurden in Abschnitt 3.7 bereits vorgestellt. Auch wenn die Arbeitsbelastung durch eine Randomisierung der Verkehrslast Vorteile bei der Schaffung gleicher Randbedingungen zu allen Konfigurationen bietet, so erhöht die Maßnahme die Varianz der Aktivzeit und Reaktionszeit in einem Maße, die kleinere Effekte verdeckt. Die daraus resultierende geringe Effektstärke kann vermutlich als erste Ursache genannt werden, die Grenzwertgrößen nicht eindeutig bestimmten zu können. Dennoch scheint das konzeptionelle Vorgehen sinnvoll, mit Hilfe der Baseline eine Kalibrierung erfolgreich vornehmen zu können.

Im Folgenden werden die ursprünglichen Anwendungsszenarien von AccSis und die entwickelten Modelle unter Berücksichtigung der Erkenntnisse des Experiments diskutiert.

4.1.1 Die Effektivität des Verfahrens und mögliche Anwendungsszenarien

Aus den Erkenntnissen der experimentellen Studie können hinsichtlich der Unterstützung der für die in Abschnitt 1.4.4 definierten Anwendungsszenarien folgende Schlüsse gezogen werden:

- *Verifikation von erwarteten Risikoschwerpunkten*: Eine effektive Unterstützung für dieses Szenario erscheint möglich, da Effekte des induzierten Zeitdrucks bereits anhand von Wahrnehmungs- und Entscheidungsschwierigkeiten des Operateurs beobachtet werden können. Die entsprechenden ursächlichen Ereignisse, wie z.B. menschliche Fehler, treten i.d.R. häufiger in Erscheinung als die sicherheitsrelevanten Ereignisse selbst. Ein mögliches Induktionsverfahren würde eine stufenartige Stressvariation umfassen. Der Zeitdruck wäre dabei schrittweise von der Basislast aus zu erhöhen bis die ersten Effekte sichtbar werden. Die schrittweise Erhöhung bis zu dem möglichen Maximum würde es erlauben, die in Erscheinung getretenen Beobachtungen der jeweiligen Stufe der Zeitdrucklast zuzuordnen. Die Beobachtung kann sowohl die Qualität der Stressreaktionen (z.B. Typus des menschlichen Fehlers) als auch deren Eintrittshäufigkeit umfassen. Eine Verifikation wird möglich anhand der folgenden Methoden:
 - i. Ein quantitativer Vergleich der Ergebnisse mit einem Referenzsystem, für das die Beobachtungen eines äquivalenten Experiments vorliegen. Hier könnte eine geeignete Hypothese definiert werden, die eine mögliche Änderung der Häufigkeit der Ursachen unterstellt. Das Ergebnis des Hypothesentests könnte somit z.B. die Effektivität einer risikomindernden Maßnahme nachweisen.
 - ii. Ein qualitativer Vergleich des Rankings der Ursacheneintrittshäufigkeiten.
- *Identifikation von unbekanntem Risikoschwerpunkten*: Nach Vorbild der Verifikation von erwarteten Risikoschwerpunkten scheint ein Beitrag durch die Anwendung des Verfahrens „Konkurrenzdruck“ zur ergebnisoffenen Identifikation von noch nicht bekannten Risikoschwerpunkten vorstellbar. Ebenso hier könnte eine stufenartige Stressvariation dazu dienen, die menschlichen Fehler und Schwächen im Konzept und

Design des Systems zu manifestieren. Voraussetzung für die Durchführung ist die Definition eines sicherheitsrelevanten Ereignisses, das mit einer hinreichenden Häufigkeit beobachtet werden kann. Die gewonnene Stichprobenmenge kann sodann einer retrospektiven Ursachenanalyse dienen, die im Rahmen des aufgezeichneten Datenmaterials und Befragungen möglich wäre. Die methodischen Schwierigkeiten bei der Wahl dieses Anwendungsszenarios sind unter anderem die Folgenden:

- i. Die Wahl von Maßnahmen zur Beobachtung und Analyse ist schwierig zu begründen, wenn keine Vermutung über mögliche Risikoschwerpunkte zulässig sein soll. Die Begründung der Wahl bestimmter Aufzeichnungsmittel und –menge oder bestimmter Inhalte der Probandenbefragung sowie Auswertung unterstellt implizit die Existenz einer Hypothese, die es zu verifizieren gilt. Dieser Widerspruch offenbart, dass die rein ergebnisoffene (nur induktive) Ermittlung von Ursachen und die Frage um die Wahl der Beobachtungsmittel in einem Konflikt zueinander stehen. Die Lösung kann demnach nur in einem wechselseitigen Zusammenspiel der Vermutungsbildung und der anschließenden Verifikation liegen. Die Begründung der initialen Maßnahmen sollte dann die folgenden Argumente berücksichtigen:
 - a. Phase der Entwicklung (z.B. HMI Design oder Verfahren zur Entscheidungsfindung).
 - b. Entsprechend den Ergebnissen der Ursachenanalyse können die Ursachenklassen weiter dekomponiert werden.

Im einfachsten Fall können ergebnisoffene Befragungen der Probanden durchgeführt werden, die dann allerdings einen professionellen Hintergrund haben sollten, um das Erlebte der Simulation nach realen Maßstäben bewerten zu können.
 - ii. Die Verifikation einer möglichen Ursache hinsichtlich der Kausalität kann aufgrund des gewählten Untersuchungsansatzes nicht gelingen.
- *Bestimmung eines unbedingten Risikos:* Die Bestimmung des Risikos anhand der beobachteten Eintrittshäufigkeit des gewählten sicherheitsrelevanten Ereignisses scheint eine vielversprechende Anwendungsmöglichkeit von AccSis zu sein. Es wird von den folgenden Voraussetzungen ausgegangen:
- i. Der Eintritt jedes sicherheitsrelevanten Ereignisses wird auf Gültigkeit hinsichtlich möglicher Simulationsartefakte oder Künstlichkeit geprüft.
 - ii. Die Definition einer Grundgesamtheit für den Eintritt eines sicherheitsrelevanten Ereignisses ist zulässig. Nach Vorbild der Experimentalstudie kann dies z.B. die Startfreigabe sein, die eine Gelegenheit für den Eintritt einer Runway Incursion bildet.
 - iii. Die maximal mögliche Intensivierung des Induktionsverfahrens erlaubt in dem gegebenen Zeitraum die Erzeugung von sicherheitsrelevanten Ereignissen in einem Maße, dass eine maximal zulässige Irrtumswahrscheinlichkeit erfüllt wird.

Dieses Anwendungsszenario zeichnet in besonderem Maße aus, keine Kenntnis über das zu erwartenden Risiko vorzusetzen. Dies führt insbesondere bei dem dritten Punkt zu dem Problem, keine Aussagen über die geforderte Intensivierung und den

Zeitraum treffen zu können. Falls es möglich ist, die Anzahl der Gelegenheiten für das sicherheitsrelevante Ereignis zu zählen, kann das Vertrauensintervall nach Vorbild der Hypothesenprüfung auf Basis der Binomialverteilung in Abschnitt 3.9.5 eine Auskunft über die erreichte Qualität liefern.

- *Verifikation einer Zielsicherheit:* Die stressbedingte Intensivierung der Menschfehlerrate um den Faktor 10 (Swain & Guttman, 1983) und die in Kapitel 3 vorgestellte experimentelle Studie liefern keinen Anhaltspunkt, welche maximale Intensivierung der Unfallwahrscheinlichkeit durch „Konkurrenzdruck“ zu erwarten ist. Allerdings lässt die Zielwahrscheinlichkeiten von $1,55e-8$ pro Flugstunde die Geringfügigkeit des erzielbaren Beschleunigungseffektes vermuten. Eine Verifikation von unbedingten Zielsicherheiten in diesen Größenordnungen erscheint mit AccSis somit nicht denkbar. Falls dennoch die Anwendung dieses Szenarios gewählt wird, so erscheinen zwei Maßnahmen praktikabel, um die Irrtumswahrscheinlichkeit zu mindern:
 - i. Eine sehr effektive Maßnahme ist, ein Vorfalleignis statt des Unfallereignisses für die retrospektive Risikoanalyse zu wählen. Dies Vorgehen wurde ebenso in der Pilotstudie gewählt. Wichtig ist hierbei allerdings, dass ein kausaler Zusammenhang zum Unfallereignis bekannt ist und in dem Betriebskonzept des neuartigen Systems angenommen werden darf. An dieser Stelle liefert der Fehlerbaum des Accident-Incident-Models (SESAR, 2012) geeignete Ereignisse als alternative Risikometrik. Bei der Untersuchung der Runway Collision ist z.B. die Wahl der Runway Incursion zulässig, die gemäß des Accident-Incident-Models eine Eintrittshäufigkeit von $3,0e-5$ pro Bewegung kennzeichnet.
 - ii. Eine Nachbefragung von Expertenprobanden könnte zusätzliche Aussagen zu kritischen Situationen und Vorfälle in situ liefern. Ein Risiko ließe sich ermitteln, wenn alternative Verläufe der jeweiligen Situation zu einem hypothetischen Unfall geführt hätten. Es ist hierbei entscheidend, die Bedingungen des alternativen Verlaufes und Wahrscheinlichkeiten dafür zu ermitteln. Auch wenn es sich hier um subjektive Aussagen des Experten handelt, kann die erlebnisnahe Begegnung mit der simulierten Situation in persona eine realistische Einschätzung liefern, welche situativen Faktoren für eine weitere Eskalation gefehlt hätten. Solange die situativen Faktoren nicht durch die Neuheit des Systems betroffen sind, ist das Betriebswissen des Experten gültig.

Aufgrund der Nutzung des retrospektiven Ansatzes ist es denkbar, die Anwendungsszenarien vereint in einem Versuchsaufbau zu verwenden. Die Kombination von offener (induktiver) Ermittlung und der Verifikation von Risikoschwerpunkten kann zu einer wirksamen Wechselwirkung führen, in der das Wissen über die Unfallentstehung eine iterative Weiterentwicklung erfährt. Die schrittweise Erhöhung der Zeitdrucklasten wäre zudem auch bei der Verifikation von Zielsicherheiten sinnvoll, da die angefahrenen Lasten als Stützstellen im zeitdrucksensitiven Bereich dienen können. Die Methode AccSys könnte mit Hilfe dieser Anwendungsszenarien Sicherheitsnachweise (Safety Evidence) zu allen Phasen des Safety-Case-Lebenszyklus eines soziotechnischen Systems mit Ausnahme der Inbetriebnahme erbringen.

Alle Anwendungsszenarien unterliegen dem Problem, dass die Bedingungen der Simulation das Auftreten eines gewählten sicherheitsrelevanten Ereignisses aufgrund fehlender externer Validität möglicherweise nicht zulassen. Dieses Problem wurde in Abschnitt 1.4.3.1 ausführlich vorgestellt und betrifft alle Beobachtungen und Rückschlüsse, die am Simulator vorgenommen werden. Ebenso kann dies logischerweise den Fall betreffen, dass das gewählte sicherheitsrelevante Ereignis aufgrund fehlerhafter Bedingungen des Simulators auftritt. Auch mit der technischen Weiterentwicklung der nächsten Jahre können die Probleme der Simulation vom Prinzip her allenfalls gemindert aber nicht eliminiert werden. Aus diesem Grund sollten Experten die Simulationsszenarien und die Bedingungen begutachten, unter denen die Risikoanalyse durchgeführt wird. Eine Verringerung der Künstlichkeit nach Berücksichtigung der Expertenaussagen erscheint möglich. Es ist aufgrund dieses konzeptionellen Schwachpunktes des Simulators zu empfehlen, AccSis nicht als alleiniges Mittel der Risikoanalyse zu wählen. Die Qualität von AccSis ist auch weiterhin auf die Mithilfe konventioneller Methoden angewiesen. Die gegenseitige Unterstützung führt dann zu einer Gesamtsteigerung der Qualität der Risikoanalyse.

AccSis ist konzeptionell anpassbar an die Wahl eines alternativen Stressors und dazugehörigem Induktionsverfahren. Dies ist insbesondere sinnvoll, wenn der Zeitdruck durch Einschränkungen des Anwendungskontextes als Stressor nicht geeignet erscheint. Es ließen sich hier Stressoren wie z.B. die Leistungsunterschiede von Verkehrsbewegungen pro Zeitraum nennen. Bei der Wahl eines Stressors sollten allgemein die folgenden Kriterien mit berücksichtigt werden:

- Steuerbarkeit der Intensität
- Kalibrierbarkeit der Intensität
- maximale erzeugte Intensität
- mögliche Aneignung bzw. Erlernen von Kompensationsstrategien durch den Probanden wie z.B. die Abwägung von Schnelligkeit und Genauigkeit.

Bei der Entwicklung eines geeigneten Induktionsverfahrens sollten diese kontextspezifischen Argumente einbezogen werden:

- Effektivität des Verfahrens zur Induktion des Stresses („Welches Verhalten des Probanden belohnt das Induktionsverfahren?“),
- Beeinträchtigung der Arbeit durch das Verfahren (z.B. Störung des Arbeitsmusters und Aufmerksamkeitsverteilung),
- Kovarianz der Zeitdrucklast auf die Qualität der entstehenden Fehlertypen und
- Konflikt mit den Arbeitsverfahren.

Die experimentelle Studie hat gezeigt, wie sehr die Effektivität der Intensivierung der Unfallwahrscheinlichkeit von der gewählten Implementierung in den Versuchsaufbau abhängt. So mussten Verbesserungen an der Visualisierung der Uhr vorgenommen werden, um den Probanden die Verknappung des Vorsprungs effektiv vor Augen zu führen. Es können Funktionsmechanismen entscheidend sein, die gewünschten Anreize für den Probanden herzustellen, wie z.B. das Prinzip der Bestrafung. Die Entwicklung und die Diskussion der Implementierung des Induktionsverfahrens hatten zahlreiche alternative Möglichkeiten der Fehlerstimulation in Erwägung gezogen (z.B. Geld als Belohnung für schnelles Arbeiten). Ausschlaggeben für das Verwerfen von Implementierungsvarianten waren in den häufigsten

Fällen die Störung des Arbeitsflusses und die Beeinflussung der geltenden Verfahren. Gerade das Argument der Unantastbarkeit der Aufgabendefinition ist bei der Bewertung möglicher Gestaltungsvarianten von entscheidender Relevanz. Die Nutzung des Wettbewerbs als suggestives Mittel der Induktion von Zeitdruck hat sich an dieser Stelle positiv bewährt.

4.1.2 Rückschluss auf die Modelle zur Bestimmung des Risikos

4.1.2.1 Unfallverteilung

Die Analyse der Stressreaktionen der Pilotstudie konnte anhand der Anzahl der Runway Incursion zeigen, dass die Beschreibung der Verteilung zu einer Zeitdrucklast möglich ist. Eine Regression erscheint somit möglich für den Fall, dass mehr als zwei Stützpunkte vorliegen. Allerdings haben es die Versuchsbedingungen und die Aufgabendefinition erlaubt, durch das Erteilen der Startfreigabe die Anzahl der Gelegenheiten für eine Runway Incursion zu bestimmen. Ein modellbasierter Ansatz, wie in Abschnitt 3.9.5 anhand der Binomialverteilung gezeigt, erscheint aufgrund der simplifizierten Aufgabendefinition des Probanden vertretbar. Es wäre sinnvoll, in zukünftigen Studien zu ergründen, ob ein relevanter Unterschied oder Fehler entstände, wenn die Einzelhandlungen der Probanden dennoch als unabhängig angenommen würden.

Die dazu alternative Verteilung gründet sich auf der Anzahl sicherheitsrelevanter Ereignisse pro gewählter Zeiteinheit. Die experimentelle Studie *Single Remote Tower* (Abschnitt 1.4.2.3) bot hierzu mit 13 Häufigkeitswerten pro Versuchskonfiguration die Möglichkeit, anhand des Medians und der Quantilen zu argumentieren. Hier konnten bereits klare Trends erkannt werden. Am Beispiel der Pilotstudie in Kapitel 3 ist erkennbar, dass sich hier ebenso wie bei der Studie *Single Remote Tower* die Dauer einer Versuchsdurchführung als Zeiteinheit anbietet.

4.1.2.2 Rückschluss auf die Stress-Risiko-Relation

Ein möglicher Kandidat, der für eine Stress-Risiko-Relation Vorbild stehen kann, wurde bereits in Abschnitt 2.6.2.3 vorgestellt und diskutiert. Einen Hinweis auf Gültigkeit des qualitativen Verlaufes der Kurve konnten die Ergebnisse nicht liefern. Allerdings konnten einige Varianzen gefunden werden, die innerhalb einer Relation konzeptionell berücksichtigt werden sollten.

Die Pilotstudie konnte hierzu feststellen, dass Kandidaten bei der gleichen Zeitdrucklast beobachtbare unterschiedliche Verhaltensweisen in den gemessenen Größen aufzeigen. Der Einfluss persönlicher Merkmale auf die Bestimmung des Risikos ist messbar.

- *Risiko und Zeit*: Relevante Einflussfaktoren sind die individuelle Basisleistung und die Abwägung zwischen den Ansprüchen nach Risiko und Zeit. Forschungstätigkeiten in Zusammenhang mit der Modellierung des Abwägungsprozesses zwischen Genauigkeit und Schnelligkeit zeigen Ansätze auf, die das Verhalten der Probanden erklären können. Eine Berücksichtigung dieser Varianz durch eine geeignete Relation würde zu einem SAT-ähnlichen Konzept führen, das die Leistungsfähigkeit des Operateurs anhand von zwei Parametern beschreibt. Dieses Konzept wurde in Abschnitt 3.9.5 für die Modellbildung benutzt und konnte somit eine Erklärung für die Stressreaktion liefern. Das Regressionsmodell könnte berücksichtigen, dass eine Abhängigkeit zwischen den Parametern besteht, die im einfachsten Fall wie bei den Kandidaten A

und C linear ist. Das Regressionsmodell würde dann die beiden folgenden Parameter berücksichtigen:

- i. Individuelle Basisleistung und
 - ii. Proportion zwischen Risiko und Zeit.
- *Kovarianz der Fehlertypen*: Eine besondere Herausforderung stellt der unbekannte Effekt dar, den erzeugten Fehlertypus durch steigenden Zeitdruck systematisch mit zu variieren. Eine derartige Kovarianz kann aufgrund des wirkenden Zeitdrucks neue Typen von Fehlern hervorrufen, die zu Bedingungen des Designstresses nicht vorliegen. Diese Typenverschiebung, in der das Profil der beobachteten Fehlertypen eine Änderung bei zunehmendem Zeitdruck erfährt, konnte Bubb beobachten (Bubb & Jastrzebska-Fraczek, 1999). Der Effekt könnte eine Verstärkung unfallverursachender Verhaltensmuster verstärken, die in einem Realbetrieb keine Relevanz haben. Bei der Identifikation von Risikoschwerpunkten im hohen Lastenbereich könnten dann menschliche Fehler zu finden sein, die überproportional zeitdruckspezifisch sind. Spezifische Ausprägungen und Eskalationsmechanismen des menschlichen Fehlers zu bestimmten Lastenbereichen könnten erklären, dass die Stress-Risiko-Relation unstetig über den Arbeitsbereich des induzierten Zeitdrucks ist und aufgrund der unterschiedlichen Risikosensitivitäten zu unterschiedliche „Phasen“ der Kurve führen kann. Eine geeignete qualitative Form der Stress-Risiko-Relation könnte diese Kovarianz durch die Definition von Abschnitten berücksichtigen. Eine geeignete Einteilung wäre an den Stressreaktionen von mehreren Zeitdrucklasten mit geringerer Schrittweite und mehr Stützpunkten zu ergründen.

4.1.3 Implikationen durch Novizen- und Expertenprobanden

Die Ergebnisse verdeutlichen einen systematischen Nachteil der Benutzung der Echtzeitsimulationen. Der Novizenproband befand sich in einer Arbeitsumgebung, die ihm keine ernsthaften betrieblichen Konsequenzen bot. Dies zeigt die Anzahl der Runway Incursion während der Durchführung der Baseline. Im realen Flugbetrieb könnten diese Entscheidungen aber das eigene oder das Leben anderer Menschen betreffen. Eben dieser Effekt könnte auch bei Verwendung von Experten für die Echtzeitsimulation zum Tragen kommen. Allerdings projiziert der Experte die möglichen Konsequenzen des Realbetriebs auf einer mentalen Ebene in den simulierten Betrieb. Dieses könnte anhand der routinierten und standardisierten Arbeitsweise von Experten zu erklären sein, wodurch er das angeeignete Bewusstsein und den Erfahrungsschatz über mögliche Konsequenzen der Handlungen auch innerhalb einer Echtzeitsimulation wohl nicht ablegt. Wie in der Einleitung 1.1 erläutert, können die geänderten Randbedingungen am Arbeitsplatz, die mit der Einführung eines neuartigen soziotechnischen Systems einhergehen, diese Routine affektieren. In welchem Maße dies auch das Bewusstsein über mögliche Konsequenzen der eigenen Handlungen betrifft, ist in weiterführenden Studien zu klären.

Schlussfolgernd betrifft dieses Problem, das Fehlen eines solchen Bewusstseins nicht kompensieren zu können, verstärkt Novizen. Die Verwendung von Novizen ist bei derartigen Untersuchungen möglich und üblich, da die höhere Verfügbarkeit eine höhere Stichprobenanzahl zum Ergebnis hat. Grundsätzliche Verhaltensmuster des Menschen, gleich ob Novize oder Experte, sind ähnlich und unterscheiden sich bei neuartigen soziotechnischen

Systemen nicht wesentlich (Vu, Minakata, & Nguyen, 2009). Dies setzt ein Training voraus, das den Novizen für die Aufgaben im ausreichenden Maße für die Aufgabe qualifiziert. Dennoch sind seltene Situationen zu beachten, in denen der Novize gegenüber dem Experten wenig oder kein Erfahrungswissen aufweist und so eine erhöhte Beanspruchung erfährt, die so beim Experten nicht zu beobachten wäre. Die Qualifikation, Ausbildung und die Erfahrung von Experten bewirkt in einem sicherheitsrelevanten Betrieb ein höheres Bewusstsein über die Konsequenzen der Handlungen. Dies lässt sich anhand der Qualität der Schätzung über die weitere Entwicklung des Betriebs feststellen:

„For comprehension questions, students were less accurate on questions about future events compared to present events, but ATCs showed higher accuracy for future than present events.” (Vu, Minakata, & Nguyen, 2009, S. 873)

Gerade derartige Unterschiede lassen den Schluss zu, die empirische Studie mit Hilfe von Experten zu wiederholen, die unter Einbezug eines höheren Betriebskontextes arbeiten. Die Reproduktion der Ergebnisse hätte somit externe Validität hinsichtlich der verwendeten Population. Es ist zu vermuten, dass sowohl die Routine als auch die Erfahrung von kritischen Grenzfällen ein hohes Systemverständnis erzeugen und eine höhere Robustheit gegenüber Stresseinwirkung bewirken. Der Zusammenhang zwischen Stress, Risiko und Zeitfehler wird durch diese Robustheit nicht grundsätzlich aufgehoben, aber nach Leistungsfähigkeiten anders skaliert und den Arbeitspunkt vermutlich zu Gunsten des Zeitfehlers verschieben. So dürfte der Proband C aufgrund seiner Proportionen zwischen RI-Rate und Zeitfehlern dem realen Vorbild vermutlich am nächsten kommen.

Die Ergebnisse haben bei Proband B den Effekt gezeigt, sich für ein höheres Risiko zu entscheiden, um Zeit zu gewinnen und den Vorsprung gegenüber dem Referenzoperator auszubauen. Die Entscheidung des Probanden hierfür kann als Folge des *Scrambled Mode* auftreten, aber auch eine strategische Entscheidung darstellen, sich dem Leistungsvergleich durch *ComPerf* so zu entziehen. Die Ursache lässt sich anhand des Bestrafungsmechanismus erklären, der Zeitfehler bestraft und RI unbestraft belässt. So ist es als Konsequenz plausibel, den Bestrafungsmechanismus des *ComPerf*-Konzeptes derart zu erweitern, den Eintritt einer Runway Incursion ebenso ausgleichend zu sanktionieren. So würde die Sicherheit durch den Operator eine höhere Arbeitspriorität erhalten, die das Verhältnis zwischen Zeit und Qualität neu gewichtet.

4.2 Die Anwendung auf andere Arbeitspositionen des Flugbetriebs

Aufgrund des Zuschnitts des Konzepts AccSys auf soziotechnische Systeme und der Modifikation der Simulationsbedingungen durch das Einwirken eines Stressors ist jedes System geeignet, ein möglicher Gegenstand von Untersuchung mit Hilfe von AccSys zu sein. Relevant ist hierbei, dass die Echtzeitsimulation die Risikoanalyse unterstützt. Dies betrifft im Wesentlichen den Unfallhergang, der nur eine geringe Beeinträchtigung durch die fehlende Gültigkeit der Simulationsbedingungen erlaubt. Als mögliche Anwendungsfelder könnten prinzipiell alle Arbeitspositionen des Flugbetriebs dienen. Tabelle 4 (Seite 77) zeigt beispielhaft, welche Positionen in die engere Auswahl gehören könnten.

Bei der Wahl der Position ist die Entwicklung einer geeigneten Implementierungslösung des Induktionsverfahrens die herausfordernde Problemstellung für

eine erfolgreiche Anwendung. Es werden im Folgenden zwei Positionen exemplarisch diskutiert und eine entsprechende Implementierungsvariante vorgestellt.

- *Luftfahrzeugführer*: Die Position umfasst neben anderen das Kontrollieren und das Navigieren des Luftfahrzeugs sowie die Kommunikation mit der Flugsicherung. Sie weist mit diesem Hintergrund eine extrem vielschichtige Aufgabendefinition aus. Die Entwicklung eines geeigneten Zeitdruckinduktionsverfahrens ist aufgrund der vermutlich höheren Komplexität der Entscheidungsfindung schwieriger als bei den Positionen der Flugsicherung. Ebenso ist die A-priori-Eingrenzung aller möglichen Situationen, die bei der Einführung neuer Systeme eine Rolle spielen könnten, schwierig zu bestimmen oder gar zu simulieren. Dies betrifft insbesondere die unterschiedlichen klimatischen Bedingungen der Erde als auch die unterschiedlichen Sprachen und Kulturen. Die einfachsten Stressoren sind beispielhaft vorgestellt.
 - i. Ein unkontrollierbarer Brand an Bord während des Fluges gehört zu den kritischsten Situationen des Luftfahrzeugführers bzw. der Crew. Die Nachricht über einen Brand signalisiert die höchste Zeitdringlichkeit, zu landen. Eine derartige Situation kann in einem beliebigen Moment der Simulation unterstellt werden. Ebenso können Notszenarien mit begrenzten Treibstoff eine Zeitdringlichkeit suggerieren. Es ist dabei zu beachten, dass solche Luftnotlagen durch die Flugsicherung als sogenannte *Distress Phase* kategorisiert werden. Ihnen wird gemäß ICAO PANS-ATM Doc. 4444, §15.1.3 Priorität gegeben (ICAO, 2007b). Ein derartiger Stressor erzeugt eine Notfallsituation und weicht aus diesem Grund erheblich von dem täglich üblichen Betrieb ab. Ebenso von Nachteil ist die vermutlich wenig kontrollierbare Intensität der Stressinduktion.
 - ii. Die Vorgaben von Zeiten für die späteste Ankunft am Zielflughafen aufgrund des Betriebsschlusses des Flughafens oder das späteste Überfliegen eines Wegpunktes sind einfache Maßnahmen der Fristsetzung. Sie bieten eine einfach zu dosierende Zeitdringlichkeit und können einfach in der Implementierung des Experimentaldesigns umgesetzt werden.
 - iii. Die Unterstellung eines konkurrierenden „Verfolgers“, der die gleiche Strecke fliegt, suggeriert dem Luftfahrzeugführer/ der Crew die Situation eines Wettrennens. Ein derart gestaltetes Verfahren kommt dem Induktionsverfahren „Konkurrenzdruck“ konzeptionell am nächsten und arbeitet ohne direkten Eingriff in die Arbeit des Operateurs.

Diese drei Vorschläge offenbaren, dass es keine Universallösung für eine effektive Vermittlung von Zeitdruck gibt. Je komplexer die Aufgaben sind, desto situationsspezifischer sollten die Untersuchungen grundsätzlich zugeschnitten werden. Im Falle des Luftfahrzeugführers können diese nach Flugphasen eingeteilt werden, in welche nach z.B. Lande- oder Startvorgängen als abgegrenzte Untersuchungsfelder unterscheiden werden kann.

- *Streckenkontrolle*: Die Aufgabendefinition des Fluglotsen hat vermutlich eine geringere Komplexität als die der Flugplatzkontrolle. Dies ist vornehmlich durch die Außensicht des Towers begründet, die ein intensives und wichtiges Mittel zur Situationswahrnehmung darstellt. Das Verfahren „Konkurrenzdruck“ kann im einfachsten Fall direkt übertragen werden.

4.3 Ein Erklärungsansatz mit dem Contextual Control Modell

Wie bereits schon im vorhergehenden Abschnitt erwähnt, wiesen die Stressreaktionen der Probanden A und C ein Zeitdrucklasten-übergreifendes Verhältnis von RI-Rate zu Zeitfehlern auf. Proband B hingegen zeigte insbesondere bei der zweiten Laststufe eine zunehmende Risikobereitschaft. Dieser Unterschied zu den beiden anderen Kandidaten ist möglicherweise anhand eines Wechsels der Arbeitsstrategie zu erklären, der durch den steigenden Zeitdruck ausgelöst wurde. Ein möglicher Ansatz zur Erklärung liefert das *Contextual Control Model* (CoCoM) nach Hollnagel (Hollnagel, 1993). Dieses Modell beschreibt einen Zusammenhang zwischen Stresseinwirkung und dem Maß an Kontrolle über eine Situation bei Ausübung einer Aufgabe. Das Modell unterscheidet hierzu nach vier Modi:

- *Strategic Mode*: Der Operateur hat einen Handlungshorizont und kann höhere Betriebsziele berücksichtigen. Dieser Modus liefert eine effiziente und robuste Arbeitsleistung und stellt das Ideal dar.
- *Tactical Mode*: Der Operateur bedient Ziele, die über die gegenwärtig verlangten hinausgehen. Die möglichen Handlungen sind aber an den unmittelbaren Kontext gebunden.
- *Opportunistic Mode*: Der Operateur handelt nach dem unmittelbar Verlangten und nach den wichtigsten Merkmalen des Kontexts.
- *Scrambled Mode*: Der Operateur handelt zufällig und unvorhersehbar.

Hollnagel identifizierte die zur Verfügung stehende Zeit neben anderen Stressfaktoren als einen entscheidenden Faktor für das Auslösen eines Moduswechsels. Bei dem Vergleich der Stressreaktionen der Probanden mit den CoCoM-Modi kann das Verhalten von Proband B gegenüber dem von Probanden A und C gedeutet werden. Es könnte sich hier um einen reduzierten Handlungshorizont handeln, der stärker auf das unmittelbare Betriebsgeschehen gerichtet ist. Das unterschiedliche Leistungsniveau erklärt sich somit anhand der Planungsarbeiten durch Proband A und C, die Verkehrsbewegungen über einen längeren Zeitraum bei gleichbleibender Gewichtung zwischen Schnelligkeit und Risikobereitschaft sequenzieren zu können. Dieses Verhalten hat Übereinstimmungen mit dem *Strategic Mode* und dem *Tactical Mode*. Der Proband B erfährt durch die zweite Laststufe einen Wechsel: Die Risikobereitschaft steigt zu Gunsten der Schnelligkeit. Der Proband kann also einen Moduswechsel vollzogen haben, der durch das unmittelbar Verlangte des Betriebsgeschehens geleitet ist (*Opportunistic Mode*) oder gar zufällig (*Scrambled Mode*).

Diese Interpretation lässt einen grundlegenden Zusammenhang zwischen der Risikobereitschaft, der Schnelligkeit und der resultierenden Gesamtleistung bei variierendem Zeitdruck erkennen. Dies wird insbesondere durch die empirischen Studien der Wechselwirkung zwischen Schnelligkeit und Genauigkeit gestützt, nach dem die Genauigkeit gemäß des *Fast Guess Modell* (Ollman, 1966) bei abnehmender Zeit und ab einem kritischen Grenzbereich einen dramatischen Einbruch erfährt. Proband B hätte demnach den Übergang von einem kontrollierten Modus, oder auch *Stimulus Controlled Response*-Modus, zu einem

Bereich, der dem sogenannten *Fast Guess*²⁰ zuzuordnen ist, vollzogen. Ebenso würde dieses Verhalten mit der Stress-Risiko-Relation übereinstimmen, die bei einer zu hohen Zeitdrucklast eine erhebliche Herabsetzung der menschlichen Leistungsfähigkeit und somit eine Intensivierung der Unfallwahrscheinlichkeit vorhersagt.

4.4 Ein Modell zur Beschreibung von Risikobereitschaft und Schnelligkeit

Es wurden Beispiele verwendeter SAT-Kurven auf Basis des sequentiellen Sample-Modells in Abschnitt 2.5.3.1 vorgestellt. Ein Anteil der Varianz des Verhaltens lässt sich anhand eines derartigen Modells aufgrund des induzierten Zeitdrucks beschreiben und für die Verifikation des Risikos verwenden. Zwei Bedingungen werden für eine gültige Übertragung von SAT zur Stress-Risiko-Relation unterstellt:

- Die Genauigkeit (eng. accuracy) repräsentiert die Qualität der Arbeit und korreliert mit einer Risikobereitschaft des Operateurs, durch sein Verhalten eine minimal geforderte Qualität zu unterschreiten und somit Fehler zu provozieren.
- Die Reaktionszeit korreliert mit der Anzahl an Zeitfehlern. Je höher die Anzahl an Zeitfehler, desto weniger Reaktionszeit weist die Arbeitsgeschwindigkeit des Operateurs auf.

Anhand der Stressreaktionen der drei Probanden der Pilotstudie erscheint die Entwicklung einer Kurve für den vorliegenden Untersuchungskontext möglich. Die Kurve beschreibt einen spezifischen Verlauf des Zeitfehlers und der Risikobereitschaft, wenn der induzierte Zeitdruck beständig steigt. Die Risikobereitschaft ist dabei das durch den Operateur gewählte Maß des akzeptierten Risikos zu einer spezifischen Zeitdrucklast. Im Folgenden werden mögliche qualitative Merkmale diskutiert, denen eine SAT-Kurve durch den Arbeitsbereich des induzierten Zeitdrucks folgen könnte.

- *Stetige vs. unstetige Form*: Beide Formen der Kurve sind denkbar und wären plausibel. Es ist möglich, dass die Leistungsfähigkeit mit Hilfe eines diskontinuierlichen Zwei-Phasen-Modells (Phase Transition Model) beschrieben wird. Allerdings ist hierbei unklar, wie der Schwellenwert des Phasenwechsels zu erklären ist. Die Stress-Risiko-Relation gemäß Abschnitt 2.6.2 deutet demgegenüber auf einen kontinuierlichen Übergang hin, bei dem sich der Operateur einem zufallsbasierten Verhaltensmuster schrittweise annähert.
- *Konkave vs. sigmoide Kurve*: Ebenfalls sind auch hier beide Formen der Kurve denkbar. Die Stress-Risiko-Relation gemäß Abschnitt 2.6.2 könnte darauf hinweisen, dass die Risikobereitschaft zu höheren Zeitdrucklasten schlagartig zunimmt und keinen sigmoide Form annimmt. Bei Unterstellung dieser Form würde sich der *Scrambled Mode* gemäß CoCom schließlich durch einen rückläufigen Trend auszeichnen, die Risikobereitschaft weiter zu erhöhen. Beide Hinweise deuten auf einen konkaven Verlauf hin.
- *Grenzrisiko*: Die Übertragung des *Contextual Control Models* auf die vorliegende Beobachtung der Pilotstudie begründet die Vermutung, dass der Proband bei

²⁰ *Fast Guess*: Die Genauigkeit der Aufgabenausführung ist Zufall (Dutilh, Wagenmakers, & Visser, 2011, S. 212).

zunehmendem Zeitdruck alle Modi der Reihenfolge nach durchlaufen könnte. Die Bedienung höherer Ziele ist hierbei ein wesentliches Merkmal der Modi *Strategic Mode* und *Tactical Mode*. Zu den höheren Zielen kann dabei z.B. die Beachtung eines Grenzrisikos gehören, unter dem der Operateur bereit ist, innerhalb des Flugbetriebs zu arbeiten. Bei Erreichen dieser Grenze würde der Operateur den Betrieb einstellen oder geeignete Mittel zur Kompensationen oder zur Minderung der betrieblichen Last wählen. Dies kann z.B. die Schließung des Flughafenbetriebs als Folge sicherheitsrelevanter Wettereinflüsse sein. Die Gewährleistung der Sicherheit wäre in diesem Fall ein Anspruch, die der Operateur der Schnelligkeit überordnet. So unterscheidet die Modi *Strategic Mode* und *Tactical Mode* das Niveau des Grenzrisikos, auf das der Proband den Betrieb begrenzt. Der *Opportunistic Mode* würde demgegenüber einen derartigen Betriebsabbruch nicht einleiten und den unmittelbaren Anforderungen des Betriebskontextes nachgeben. Eine weitere Eskalation der Situation ist somit möglich.

Auf Grundlage der drei soeben diskutierten Merkmale zeigt Abbildung 52 den Entwurf einer möglichen Kurve. Durch die Inversion der Achsen wird der konkave Verlauf zu einem konvexen Verlauf gespiegelt. Die Kurve bietet die Möglichkeit, Probanden einem Kontrollmodus zuzuordnen und die Leistungsunterschiede zu erfassen oder gar unter Einbezug einer Basisleistung zu normieren. Ebenso wäre die Form der Kurve geeignet, zeitdruckvariante Typen des menschlichen Fehlers einer spezifischen Phase zuzuordnen.

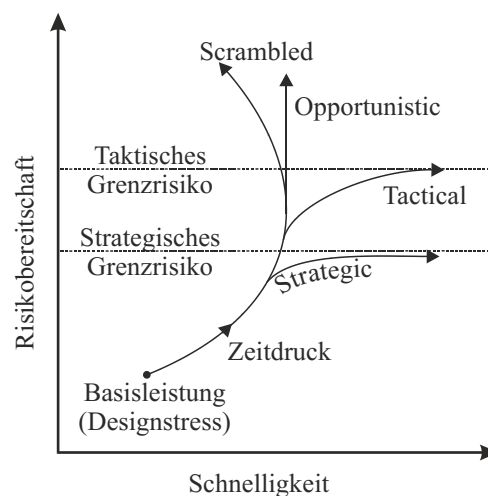


Abbildung 52: Übertragung Hollnagel CoCoM auf eine SAT-Kurve der zeitdruckabhängigen Risikobereitschaft

4.5 Mögliche Hypothesen für eine Folgestudie

Anhand der Schlussfolgerungen dieses Kapitels werden die wesentlichen Vermutungen anhand der folgenden Hypothesen vorgestellt:

- Die Stress-Risiko-Relation folgt nach Vorbild der Abbildung 52 einem konvexen Verlauf.
- Die Metrik der Leistungsfähigkeit ist zeitdrucklastsensitiv und eignet sich zur Bestimmung der Basislast t_{a0} und der dazugehörigen mittleren benötigten Zeit t_{n0} .

- Die Effektivität des Induktionsverfahrens kann gesteigert werden, wenn die Erzeugung sicherheitsrelevanter Ereignisse ebenso bestraft wird wie der Zeitfehler.
- Die Nutzung von Expertenprobanden wird das Verhältnis von Risiko und Zeitfehler aufgrund der höheren Unempfänglichkeit gegenüber Zeitdruck zugunsten der Zeitfehler verschieben.
- Die Wahl der einzelnen Entscheidungssituationen als statistische Einheit der Grundgesamtheit führt zu einem vernachlässigbaren Fehler bei der Ermittlung der Unfallverteilung gegenüber der Wahl einer Zeiteinheit.

4.6 Schlusswort

Die in dieser Arbeit entwickelte Methodik und die vollzogene empirische Studie sollen für die zukünftige Methodenentwicklung einen Beitrag leisten, die gegenwärtig praktizierte expertenbasierte Risikoanalyse mit den empirischen Untersuchungsmethoden der Anthropologie zu verbinden. Dies folgt einem Aufruf, der am OECD Halden Reactor Project verfasst wurde (Boring & Bye, 2008), angewandte Sicherheitswissenschaften und empirische Forschung verstärkt zu kombinieren.

Es wurden verschiedenste Methoden und die damit verbundenen Modelle unterschiedlicher wissenschaftlicher Disziplinen für den vorliegenden Problemfall dieser Arbeit und ihre Lösung eingebunden.

- *Accelerated Life Testing* und
- *Speed vs. Accuracy*.

Im Bereich des Luftverkehrsmanagements gab es einen „Weckruf“, der sich auf die Probleme der anstehenden „Deployment“-Phase von SESAR berief:

„SESAR could be the most complex ‘IT + human agent-based’ safety-critical system in the world. SESAR must adopt the best practices when making decisions about safety-critical changes” (Brooker, 2010, S. 842).

So werden soziotechnische Systeme in Betrieb gesetzt, wie z.B. A-SMGCS und Multi Remote Tower, deren Änderungen der Randbedingungen für den Flugbetrieb sehr stark ausfallen können und an denen etablierte Methoden an ihre Grenzen stoßen. Auch Brooker plädiert für den verstärkten Einbezug empirischer Untersuchungsmittel.

Bevor das Konzept *Accelerated Risk Analysis* Anwendung findet, sollten die o.a. Punkte fortführend untersucht werden, um z.B. die entstehende Varianz der Zielgrößen anhand einer größeren Gruppe von Probanden zu bestimmen. Als Hauptprobleme der Methode können die Folgenden genannt werden:

- Bestimmung der Grenzwerte des Designstresses,
- unzureichende Erklärungen und Berücksichtigung von Individualvarianzen in den Stressreaktionen und
- fehlende Auslegungskriterien für die Implementierung des Induktionsverfahrens, die den Anspruch der Allgemeingültigkeit für sicherheitsrelevante Arbeitspositionen erfüllen.

Weitere Forschungsaktivitäten zu den soeben genannten Punkten sollten verstärkt auf andere Arbeitsbereiche ausgedehnt werden, wie z.B. auf die Arbeit im Cockpit, die sich in ihren inhaltlichen Herausforderungen von den Tätigkeiten des Fluglotsen qualitativ unterscheidet. So können hier z.B. die Tätigkeiten zur Kontrolle der Bordsysteme und der manuellen Führung des Luftfahrzeugs während des Starts und der Landung genannt werden, die eine enorme Aufgabenvielfalt der Pilotenarbeit umfassen. Über das Cockpit hinaus wären auch Multi-Operateur-Situationen denkbar, in denen das Prinzip auf einen Komplettbetrieb ausgedehnt wird. Der Anwendungsbereich im Training und in der Qualifikation von Personal im Flugbetrieb könnten die charakteristischen Stressreaktionen zur Einschätzung des Leistungsvermögens verwenden und eine frühe Prognose über den Kandidaten oder Operateur erlauben, welche Kontrolle er über stressreiche Situationen erlangen kann.

Eine visionäre Idee stellt die Verwendung von echten Menschen als Referenzoperateure dar, die eine Wettbewerbssituation unter gleichen Anfangs- und Randbedingungen gestalten können. Die soeben aufgeführten Punkte lassen erahnen, welche methodischen Spielräume zur Gewinnung belastbarer Daten für die Risikoanalyse ausgeschöpft werden können, wenn das Konzept der simulationsgestützten Risikoanalyse sorgsam und schrittweise ausgebaut wird. Alternative Stressoren können in Erwägung gezogen werden, um den jeweiligen spezifischen Anforderungen des Einsatzes gerecht zu werden. Das grundlegende Lösungsprinzip von *Accelerated Risk Analysis*, die Arbeitsbedingungen der Echtzeitsimulation kontrolliert zu modifizieren, kann hierbei als methodischer Rahmen verstanden werden und in der praktischen Implementierung modular nach Anforderungen erweitert werden.

Abkürzungsverzeichnis

<i>Bezeichnung</i>	<i>Definition</i>
AccSis	Accelerated Risk Analysis
ALT	Accelerated Life Testing
CFIT	Controlled Flight Into Terrain
ComPerf	Konkurrenzdruck (Competitive Performance)
FAF	Final Approach Fix
FIUUG	Flugunfall-Untersuchungs-Gesetz
HEP	Menschfehlerwahrscheinlichkeit (Human Error Probability)
ICAO	International Civil Aviation Organization (Internationale Zivilluftfahrtorganisation)
IRP	Integrated Risk Picture
IWS	Integrated Workload Scale
LFZ	Luftfahrzeug
NTSB	National Transportation Safety Board (US-amerikanische Verkehrsbehörde)
P-Wert	P-Wert des jeweils durchgeführten Signifikanztests
RI	Runway Incursion
RI-Rate	Häufigkeit der Runway Incursion pro Startfreigabe
RWY	Piste (Runway)
TL	Runway Incursion, rollend - landend
TS	Runway Incursion, rollend – startend
LS	Runway Incursion, landend – startend
SAT	Speed-Accuracy-Tradeoff
SS	Runway Incursion, startend – startend

Formelzeichenverzeichnis

Zeichen	Äq. Zeichen Anhang I.2	Bezeichnung	Formelnr.	Definiert in	
				Kapitel/Absch.	Seite
α		Irrtumswahrscheinlichkeit (false positive)	2.4	2.3.1	48
a		Intensivierungsfaktor	2.13	2.4.1	52
A		Unfallereignis		2.1.1	44
γ		Relative Intensivierung: Effektive Beschleunigung der Konvergenz der Irrtumswahrscheinlichkeit α gegenüber dem Designstress	2.16	2.4.1	53
ε		Schwellenwert (hier max. tolerierbare Wahrscheinlichkeit)	2.5	2.3.2	48
λ		Parameter für die Dichteverteilung des X_n - Histogramms einer Anzahl von Simulationsläufen		2.6.1	64
k		Zeitintervall des Verkehrsgenerators mit der zeitlichen Dauer T	3.10	3.2.5.4	91
m		Zur Zeit t_{sim} vorliegendes Intervall der Menge k	3.13	3.2.5.4	92
n		In Kapitel 3: Anzahl an Zeiteinheiten pro Simulationslauf		2.3.3	49
		In Kapitel 4: Anzahl an Verkehrsbewegungen	3.11	3.2.5.4	92
$n_{\text{abgeschlossen}}$		Anzahl der zur Zeit t abgeschlossenen Verkehrsbewegungen		3.2.5.4	92
n_{aktiv}		Anzahl aktiver Verkehrsbewegungen	3.12	3.2.5.4	92
$n_{\text{generiert}}$		Anzahl der zur Zeit t generierten Verkehrsbewegungen	3.13	3.2.5.4	92
n_k		Anzahl an Verkehrsbewegungen im Intervall k	3.10	3.2.5.4	91
n_{rm}	nRo	Anzahl der erfolgreich abgeschlossenen Rollwegmodifikationen		3.2.6.2	100
$n_{\text{vorsprung}}$		Vorsprung der Anzahl abgeschlossener Verkehrsbewegungen des menschlichen Operateurs	3.15	3.2.5.4	96

Zeichen	Äq. Zeichen Anhang I.2	Bezeichnung	Formelnr.	Definiert in	
				Kapitel/Absch.	Seite
		gegenüber dem Fluglotsenagenten			
P		Arbeitsleistung	2.19	2.5.2	56
$P_{\text{konfliktfrei}}$		Maß der Leistungsfähigkeit zur konfliktfreien Freigabeerteilung über einen Zeitraum t		3.2.5.2	90
$P_{P\&S}$		Effizienz bzw. Leistungsfähigkeit der Flugplatzkontrolle	3.7	3.2.5.2	90
p		Unfallwahrscheinlichkeit	2.2	2.1.2	46
\hat{p}		Intensivierte Unfallwahrscheinlichkeit	2.13	2.4.1	52
p_0		Zielsicherheit	2.1	2.1.2	46
\hat{p}_0		Adaptierte Zielsicherheit	2.15	2.4.1	53
Q		Maß der Arbeitsqualität	2.19	2.5.2	56
$Q_{\text{konfliktfrei}}$		Maß der erbrachten Entscheidungsqualität zur konfliktfreien Freigabeerteilung	3.5 & 3.6	3.2.5.2	89
$Q_{P\&S}$		ein Maß, die Bewegungszeit zum Durchqueren des Zuständigkeitsbereiches einer Bewegung hinsichtlich eines möglichen Minimums optimiert zu haben	3.7	3.2.5.2	90
R^2		Bestimmtheitsmaß		3.1.4	78
RI		Anzahl an Runway Incursion		3.2.2	81
$RI\text{-Rate}$	RIrate	Anzahl an Runway Incursion pro Startvorgang		3.2.2	81
t_a		Verfügbare Zeit	2.20	2.5.3.2	58
t_{a0}		Verfügbare Zeit zu Bedingungen des Designstresses	2.21	2.5.4.3	62
t_{ak}		Aktionszeit zur Freigabeerteilung (Aufruf des Dialogfelds bis zur abschließenden Bestätigung)		3.2.6.2	100
t_{aktiv}		Aktivzeit einer Verkehrsbewegung vom Eintritt in die Simulation bis zur letzten durch den Operateur erteilten Anweisung	3.9	3.2.5.4	91
t_{bewegung}		Gesamte Bewegungszeit, bestehend aus Pushback (für		3.2.5.2	90

Zeichen	Äq. Zeichen Anhang I.2	Bezeichnung	Formelnr.	Definiert in	
				Kapitel/Absch.	Seite
		ausgehende Verkehrsbewegungen) , Rollen, Auf- oder Abrollen, Starten oder Landen und Anflug oder Abflug			
t_{freigabe}		Zeit von der Anfrage der Freigabe bis zur Erteilung	3.14		94
$t_{\text{konfliktfrei}}$		eine notwendige Zeit der Verzögerung zur Herstellung bzw. Gewährleistung eines konfliktfreien Flugbetriebs		3.2.5.2	90
t_L	t_L	Latenzzeit des Fluglotsenagenten		3.2.5.3	94
t_n		Benötigte Zeit	2.20	2.5.3.2	58
t_{n0}		Benötigte Zeit zu Bedingungen des Designstresses	2.21	2.5.4.3	62
t_r		Reaktionszeit des Operators		3.2.5.2	90
$t_{\text{Vorsprung}}$		Geschätzter zeitlicher Vorsprung des menschlichen Operators gegenüber dem Fluglotsenagenten bei angenommener Null-Leistung		3.2.5.5	96
T		Zeitliche Dauer eines Intervalls k für die Generierung einer definierten Menge an Verkehrsbewegungen	3.10	3.2.5.4	91
TB		Zeitbudget	2.20	2.5.3.2	58
\widehat{TB}	mTB	Angepasstes Zeitbudget	3.17	3.2.5.7	98
X_n		Anzahl beobachteter sicherheitsrelevanter Ereignisse in n Stichproben		2.2	47
ZF	nZF	Zeitfehler		2.5.3.2	58
		Häufigkeit beobachteter Zeitfehler	3.1	3.1.1.2	76
		Eintrittsbedingung bei $n_{\text{Vorsprung}}$ kleiner Null	3.16	3.2.5.6	97

Abbildungsverzeichnis

Abbildung 1: LFV- Remote Tower Center in Sundsvall	3
Abbildung 2: Vor- und Unfallraten der U.S. Marineluftfahrt von 1977–1992 (Shapell & Wiegmann, 1996)	7
Abbildung 3: Ursachenanteile des menschlichen Fehlers in der kommerziellen Luftfahrt	9
Abbildung 4: Typische Modelle zur Abbildung von Relationen zwischen Ursachen und dem Unfallereignis ..	12
Abbildung 5: ATM Safety Iceberg (Blom H. , et al., 2001)	15
Abbildung 6: Die Comet G-ALYU im Wassertank des Versuchsstandes	16
Abbildung 7: Der Evaluationskreislauf der Luftfahrtsicherheit.....	18
Abbildung 8: Die Falle der Überregulierung (Amalberti, 2001).....	20
Abbildung 9: Der <i>Ruggles Orientator</i> (Windsor, 1928).....	25
Abbildung 10: Prinzip der Echtzeitsimulation für die Risikoanalyse	28
Abbildung 11: Zentralperspektive (links) und gekreuzte Perspektive (rechts)	30
Abbildung 12: Häufigkeiten einer Runway Incursion.....	31
Abbildung 13: Multi Remote Tower-Arbeitsplatz	33
Abbildung 14: Identifikation von Schnittmengen in den Unfallursachen (Belcastro & Foster, 2010)	35
Abbildung 15: Illustration der Modellunsicherheit anhand der Mengenlehre	36
Abbildung 16: Gültigkeit der simulierten Arbeitsumgebung bei Anwendung der Untersuchungsansätze	40
Abbildung 17: Die Verteilung des Ausfallereignisses in Relation zur Stressintensität nach ALT	41
Abbildung 18: Geometrische Interpretation der Irrtumswahrscheinlichkeit des Experimentalfalls.....	50
Abbildung 19: Konvergenz der Irrtumswahrscheinlichkeit für fünf hypothetische Messwerte	52
Abbildung 20: Die Irrtumswahrscheinlichkeit für sechs beispielhafte Intensivierungen	54
Abbildung 21: Leistungsfähigkeit in Abhängigkeit des vorliegenden Zeitdrucks (Rastegary & Landy, 1993) ..	56
Abbildung 22: Stetige SAT-Kurve, modellbasiertes Verhalten (1 und 2).	58
Abbildung 23: Verlauf der benötigten und verfügbaren Zeit über das Zeitbudget	59
Abbildung 24: Das Wettbewerbsverfahren setzt den menschlichen Operateur und einen Referenzoperateur ..	62
Abbildung 25: Benötigte Zeit als Stressreaktion der verfügbaren Zeit	64
Abbildung 26: Konzept der stressinduzierten Intensivierung der Wahrscheinlichkeit des Unfallereignisses ..	65
Abbildung 27: Einfaches Risikomodell.....	67
Abbildung 28: Erweitertes Risikomodell	69
Abbildung 29: Menschfehlerwahrscheinlichkeit abhängig vom Zeitdruck	70
Abbildung 30: Der Surface Movement Manager	86
Abbildung 31: Auswahl an möglichen Situationen im simulierten Flugbetrieb	87
Abbildung 32: Die Aktivzeit als verbildlichte Reisezeit von der Quelle bis zur Senke	90
Abbildung 33: Wahrscheinlichkeitsdichte der Aktivzeitverteilung	93
Abbildung 34: Das Leistungsmodell verbindet die Zeit der Einzelentscheidung und Qualität	94
Abbildung 35: Prinzip des Leistungsvergleiches zwischen Fluglotsenagenten und menschlichem Operateur...96	
Abbildung 36: Visuelle Anzeige der Zeitreserve des menschlichen Operateurs als Konzeptskizze	97
Abbildung 37: Angepasstes Risikomodell.....	99
Abbildung 38: Verkehrsbelastungen der Kalibrierungsszenarien Dauerbelastung	105
Abbildung 39: Kalibrierungsverlauf der Latenzzeit des Fluglotsenagenten $t_L(t)$	106
Abbildung 40: Boxplot der Aktivzeitverteilung (links) und Häufigkeitsverteilung der Aktivzeiten (rechts)...	108
Abbildung 41: Anzahl der aktiven Verkehrsbewegungen in der Simulation über die Zeit	109
Abbildung 42: Zeitlicher Verlauf der Anzahl aktiver Verkehrsbewegungen des Kalibrierungsszenarios	112
Abbildung 43: Tagesganglinie der generierten Verkehrsmenge der Basislast mit drei Spitzendurchgängen ...	113
Abbildung 44: Zeitverlauf der Reaktionszeit t_r (gestrichelt), Latenzzeit t_L (schwarz)	115
Abbildung 45: Häufigkeiten der Zeitfehler der Probanden über die gesetzte Latenzzeit t_L	116
Abbildung 46: Verlauf der Arbeitsbelastung und des Zeitdrucks für die Probanden über die Versuchszeit ...	117
Abbildung 47: Mittlere Arbeitsbeanspruchung pro Intervall über das mittlere Zeitbudget	118
Abbildung 48: Boxplot der Verteilung der Aktivzeiten t_{aktiv}	120
Abbildung 49: Arbeitsbeanspruchung und Zeitdruck-Befragung beider Runden in der Übersicht.....	124
Abbildung 50: Runway Incursion-Rate und Zeitfehlerhäufigkeiten als lastenabhängige Betriebspunkte	125
Abbildung 51: Zusammenführung der RI-Rate und Zeitfehler-Häufigkeiten beider Runden.....	126

Tabellenverzeichnis

Tabelle 1: Die Konfigurationen als Kombination der variierten Gestaltungsvarianten.....	29
Tabelle 2: Gefahrenereignisse	32
Tabelle 3: Mögliche Fälle der Entscheidung über die Erfüllung der Zielsicherheit.....	48
Tabelle 4: Übersicht einer Auswahl an Arbeitspositionen mit direktem sicherheitsrelevantem Bezug	77
Tabelle 5: Direkter Beitrag der Flugsicherung zu Unfällen von Luftfahrzeugen in der ECAC Region 2005	81
Tabelle 6: Auszug der zu erteilenden Freigaben und Anweisungen des Platzverkehrslotsen.....	83
Tabelle 7: Nicht berücksichtigte Freigaben und Anweisungen	85
Tabelle 8: Qualitäten und Zeit der konkurrierenden Operateure als Faktoren der Ausgangsleistung	95
Tabelle 9: Gegenüberstellung von zwei vorgeschlagenen Leistungsmetriken für den Leistungsvergleich	95
Tabelle 10: Äquivalenzgrößen zur Definition des angepassten Zeitbudgets	98
Tabelle 11: Ausgabeformat des Protokolls – Detektion von Runway Incursion.....	99
Tabelle 12: Ausgabeformat des Protokolls – Benutzeraktion.....	100
Tabelle 13: Ausgabeformat des Protokolls – Zusammenfassung Verkehrsbewegungen	100
Tabelle 14: Ausgabeformat des Protokolls – Bestrafungsmaßnahme	101
Tabelle 15: 8-stufige Einteilung der Arbeitsbeanspruchung und Zeitdruck für die Befragung	101
Tabelle 16: Baseline und Zeitdrucklastszenarien.....	102
Tabelle 17: Relaxationszeiten.....	108
Tabelle 18: Versuchskonfigurationen pro Proband.....	111
Tabelle 19: Regressionsanalyse des Zusammenhangs zwischen Häufigkeit der Zeitfehler	117
Tabelle 20: Bestimmtheitsmaße der linearen Regression von t_L zu Arbeitsbeanspruchung und Zeitdruck.....	118
Tabelle 21: Gewählte Parameter für die Laststufen 1 und 2.....	119
Tabelle 22: P-Werte des einseitigen Mann-Whitney-U-Tests auf Gleichheit oder Vergrößerung	123
Tabelle 23: Bestimmtheitsmaße der Zusammenhänge Latenzzeit t_L , <i>RI-Rate</i> und Zeitfehler <i>ZF</i>	125

Literaturverzeichnis

- ACI. (17. März 2014). *Worldwide Airport Traffic Report Summary*. Von <http://www.aci.aero/Data-Centre/Monthly-Traffic-Data/Worldwide-Airport-Traffic-Summary> abgerufen
- ALPA. (März 2007). Runway Incursions - A Call for Action. *ALPA White Paper*.
- Amalberti, R. (2001). The paradoxes of almost totally safe transportation systems. *Safety Science* (37), S. 109-126.
- BEA. (2012). *Final Report - On the Accident on 1st June 2009 to the Airbus A330-203 registered F-GZCP operated by Air France flight AF447 Rio de Janeiro - Paris*. Paris, FR: Bureau d'Enquêtes et d'Analyses.
- Belcastro, C. M., & Foster, J. V. (2010). Aircraft Loss-of-Control Accident Analysis. *AIAA Guidance, Navigation, and Control Conference*. Toronto, CA.
- Blom, H., & Bakker, B. (2003). *Stochastic analysis background of accident risk assessment for Air Traffic Management*. Amsterdam, NL: NLR.
- Blom, H., Bakker, G. J., Blanker, P. J., Daams, J., Everdij, M. H., & Klompstra, M. B. (2001). Accident risk assessment for advanced air traffic management. *Progress in Astronautics and Aeronautics*, S. 463-480.
- Boring, R. L., & Bye, A. (2008). Bridging human factors and human reliability analysis. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, S. 733-737.
- Bortz, J., & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human und Sozialwissenschaftler, 4. überarb. Aufl.* Heidelberg, DE: Springer.
- Brooker, P. (2010). SESAR safety decision-making: Lessons from environmental, nuclear and defense modeling. *Safety Science*, S. 831-844.
- Bubb, H. (2005). Human Reliability: A key to improved quality in manufacturing. *Human Factors and Ergonomics in Manufacturing & Service Industries* (15.4), S. 353-368.
- Bubb, H., & Jastrzebska-Fraczek, I. (1999). Human Error Probability depending on Time Pressure and Difficulty of Sequential Tasks. *Safety and Reliability*, S. 681-686.
- Buch, M. (2007). Arbeitsorganisation als Ausgangspunkt der Arbeitsgestaltung. In E. Schäfer, M. Buch, I. Pahls, & J. Pfitzmann, *Arbeitsleben!: Arbeitsanalyse-Arbeitsgestaltung-Kompetenzanalyse-Kompetenzentwicklung* (S. 18-39). Kassel, DE: Kassel University Press.
- Cain, B. (2007). *A review of the mental workload literature*. Toronto, CA: Defence Research and Development Canada .
- Callantine, T. J. (2003). Detecting and Simulating Pilot Errors for Safety Enhancement. In *SAE Transactions* (S. 425-434). American Technical Publishers Ltd.
- Cameron, A. C. (2013). *Regression analysis of count data* (Bd. 53). Cambridge University Press.
- Condra, L. (2001). *Reliability Improvement with Design of Experiments*. New York, US: CRC Press.
- Cooper, J. C. (2005). The Poisson and Exponential Distributions. (A. P. Trust, Hrsg.) *Mathematical Spectrum*(37.3), S. 123-125.
- Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society - Series B (Methodological)*, S. 187-220.
- Dahlstrom, N., Dekker, S., Winsen, R. v., & Nyce, J. (2009). Fidelity and validity of simulator training. *Theoretical Issues in Ergonomics Science* 10.4, 305-314.
- Danaher, J. W. (1980). Human Error in ATC System Operations. *Human Factors: The Journal of the Human Factors and Ergonomics*, S. 535-545.
- de Jong, H. (2004). *Guidelines for the identification of hazards: how to make unimaginable hazards imaginable?* Amsterdam, NL: NLR.

-
- DFS. (2011). *Betriebsanweisung Flugverkehrsdienste*. Langen, DE: Deutsche Flugsicherung.
- DFS. (9. April 2014). *High safety standard and a leader in technology*. Von Annual Press Briefing at DFS in Langen: http://www.dfs.de/dfs_homepage/en/Press/Press_releases/2014/09.04.2014.-Annual_Press_Briefing_at_DFS/_abgerufen
- Djokic, J., Lorenz, B., & Fricke, H. (2008). ATC Complexity as Workload and Safety Driver . Fairfax, US: International Conference on Research in Air Transportation (ICRAT).
- Druzel, M. J., & Van Der Gaag, L. C. (2000). Building probabilistic networks: "Where do the numbers come from?". *IEEE Transactions on knowledge and data engineering*(12.4), S. 481-486.
- Dutilh, G., Wagenmakers, E., & Visser, I. (2011). A phase transition model for the speed-accuracy trade-off in response time experiments. *Cognitive Science*(35.2), S. 211-250.
- EASA. (2012). *CS FSTD - Certification Specifications for Aeroplane Flight Simulation*. Köln, DE: EASA.
- EASA. (2015). *CS 25 - Certification Specifications and Acceptable Means of Compliance for Large Aeroplanes*. Köln, DE: EASA.
- Edland, A., & Svenson, O. (1993). Judgment and Decision Making Under Time Pressure: Studies and Findings. In O. Svenson, & A. Maule, *Time Pressure and Stress in Human Judgment and Decision Making* (S. 27-38). New York, US: Plenum Press.
- Emery, F., & Trist, E. (1960). Socio-Technical Systems. In C. Churchman, & M. Verhulst, *Management sciences, models and Techniques, vol. 2* (S. 83–97). Oxford, UK: Pergamon Press.
- EUROCAE. (2010). *ED-125 - Process for Specifying Risk Classification Scheme and Deriving Safety Objectives in ATM*. Paris, FR: European Organization for Civil Aviation Equipment.
- Eurocontrol. (2001). *ESARR 4 - risk assessment and mitigation in ATM*. Brussels, BE: Eurocontrol.
- Eurocontrol. (2005). *Adaption of Workload Model by Optimisation Algorithms and Sector Capacity Assessment*. Brüssel, BE: Eurocontrol.
- Eurocontrol. (2006a). *Safety Case Development Manual*. Brussels, BE: Eurocontrol.
- Eurocontrol. (2006b). *Main Report for the: 2005/2012 Integrated Risk Picture for Air Traffic Management in Europe*. Brussels, BE: Eurocontrol.
- Eurocontrol. (2010a). *Definition of A-SMGCS Implementation Levels*. Brussels, BE: Eurocontrol.
- Eurocontrol. (2010b). *Generic Cost Benefit Analysis of A-SMGCS Levels 1 and 2*. Brussels, BE.
- Eurocontrol. (2012). *Performance Review Report*. Brussels, BE: Eurocontrol.
- European Cockpit Association. (10. 11 2014). *Remote Tower Services*. (ECA, Herausgeber) Von <https://www.eurocockpit.be/stories/20141110/remote-tower-services#position> abgerufen
- FAA. (2011). *Aviation Maintenance Technician Handbook – General*. Washington D.C., US: Federal Aviation Authority.
- Federal Aviation Administration. (2014). *NextGen Implementation Plan*. Washington DC, US: FAA.
- Fischer, U., Orasanu, J., & Montalvo, M. (1993). Efficient decision strategies on the flight deck. *7th International Symposium on Aviation Psychology*, (S. 198-203). Columbus, US.
- Fitts, P. M. (1966). Cognitive aspects of information processing: III. Set for speed versus accuracy. *Journal of Experimental Psychology, Vol 71*(6), S. 849-857.
- Freedman, J. L., & Edwards, D. R. (1988). Time pressure, task performance, and enjoyment. *The social psychology of time: New perspectives*, S. 113-133.
- Galster, S. M., Duley, J. A., Masalonis, A. J., & Parasuraman, R. (2001). Air traffic controller performance and workload under mature free flight: Conflict detection and resolution of aircraft self-separation. *The International Journal of Aviation Psychology, 11*, 71-93.
- Green, R. (1985). Stress and accidents. *Aviation, space, and environmental medicine, Vol 56*(7), S. 638-641.
- Günther, T., Hildebrandt, M., Fricke, H., & Strasser, M. (2011). Contributions of advanced taxi time calculation to airport operations efficiency. *Journal of Aerospace Operations*, S. 95-106.
-

- Hart, S. G., & Staveland, L. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human Mental Workload*, S. 139–183.
- Haslbeck, A., Kirchner, P., Schubert, E., & Bengler, K. (2014). A Flight Simulator Study to Evaluate Manual Flying Skills of Airline Pilots. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, (S. 11-15). Chicago, US.
- Heinrich, H., Peterson, D., & Roos, N. (1931). *Industrial accident prevention: A safety management approach*. New York, US: McGraw-Hill.
- Hollnagel, E. (1993). *Human reliability analysis: Context and control*. London, UK: Academic Press.
- Hollnagel, E. (1998). *Cognitive Reliability and Error Analysis Method (CREAM)*. Halden, Norway: Elsevier Science Ltd.
- ICAO. (1993). *Human Factors Digest No.7-Investigation of Human Factors in Accidents and Incidents (Circular 240)*. Montreal, CA: ICAO.
- ICAO. (2004). *Advanced Surface Movement Guidance and Control Systems (A-SMGCS) Manual - Doc. 9830*. ICAO Montreal.
- ICAO. (2005). *Rules of the air - 10th edition*. Montreal, CA: ICAO.
- ICAO. (2007a). *Manual on the Prevention of Runway Incursions*. Montreal, CA: ICAO.
- ICAO. (2007b). *PANS-ATM Air Traffic Management - Doc 4444. Manual doc 4444*. Montreal, CA: ICAO.
- ICAO. (2010). *Annex 13 - Aircraft Accident and Incident Investigation 10th Edition*. Montreal, CA: ICAO.
- ICAO. (2013). *Manual doc 9995 - Manual of Evidence-Based Training*. Montreal, CA.
- ICAO. (2013a). *Annex 14 - Aerodromes Volume 1 - Aerodrome Design and Operations 6th edition*. Montreal, CA: ICAO.
- ICAO. (2013b). *Annex 19 - Safety Management 1st Edition*. Montreal, CA: ICAO.
- ICAO. (2013c). *Safety Management Manual (SMM) Doc. 9859 3rd Edition*. Montreal, CA: ICAO.
- ICAO. (22. Februar 2016). *Runway Safety Statistics*. Von <http://www.icao.int/safety/RunwaySafety/Pages/Statistics.aspx> abgerufen
- Johnson, C. (1999). Why human error modeling has failed to help systems development. *Interacting with computers* 11.5, S. 517-524.
- Jones, G. M. (1927). Are you fit to drive an airplane? *The Science News-Letter* (12), S. 113–120.
- Jörn, B. A. (2010). *Flugzeugentwicklung am Beispiel des Airbus A380: Von der Idee zur Zulassung*. Hamburg, DE: Diplomica Verlag GmbH.
- Kerlinger, F. (1964). *Foundations of Behavioral Research - Educational and Psychological Inquiry*. New York, US: Holt, Rinehart and Winston, Inc.
- Kerstholt, J. H. (1994). The effect of time pressure on decision-making behaviour in a dynamic task environment. *Acta Psychologica* 86.1, S. 89-104.
- Kirwan, B. (1992). Human error identification in human reliability assessment. Part 1: Overview of approaches. *Applied ergonomics* 23.5, S. 299-318.
- Kirwan, B. (2007). *Technical Basis for a Human Reliability Assessment Capability for Air Traffic Safety Management*. Paris, FR: Eurocontrol.
- Koch, M. (2014). *Analyse von Methoden zur Messung von Zeitdruck anhand von Verhaltensmustern in der Flugplatzkontrolle*. Dresden, DE: Technische Universität Dresden.
- Kühn, A. (2014). *Experimentelle Bestimmung des Einflusses der Anflugstaffelung auf die Eintrittswahrscheinlichkeit von Runway Incursions mittels Stressverfahren*. Dresden, DE: Diplomarbeit, TU Dresden.
- Luxhøj, J. T., & Kauffeld, K. (2003). *Evaluating the effect of technology insertion into the national airspace system*. Von The Rutgers Scholar 5: <http://rutgersscholar.rutgers.edu/volume05/luxhoj-kauffeld/luxhoj-kauffeld.htm> abgerufen

-
- MacGregor, D. (1993). Time Pressure and Task Adaptation: Alternative Perspectives. In O. Svenson, & A. J. Maule, *Time Pressure and Stress in Human Judgment and Decision Making* (S. 73-82). New York, US: Plenum Press.
- Malhotra, D. (2010). The desire to win: The effects of competitive arousal on motivation and behavior. *Organizational Behavior and Human Decision Processes*, S. 139-146.
- Meyer, L., Schultz, M., & Fricke, H. (2011). Validation of an innovative experimental safety assessment for virtual control tower HMI designs. *Proceedings Sesar Innovation Days*. Toulouse.
- Meyer, L., Schultz, M., Schmidt-Roßleben, S., & Fricke, H. (2013). Understanding the safety-relevance of the visual cue perception at a Surface Manager HMI. *3rd SESAR Innovation Days*. Stockholm, SE.
- Meyer, L., Vogel, M., & Fricke, H. (2010). Functional Hazard Analysis of Virtual Control Towers. *Proceedings of 11th IFAC*. Valenciennes, FR.
- Ministry of Transport and Civil Aviation. (1955). *Civil Aircraft Accident - Report of the Court of Inquiry into the Accidents to Comet G-ALYP on the 10th January, 1954 and Comet G-ALYY on 8th April, 1954*. 1955: Her Majesty's Stationery Office.
- Mosleh, A., B., V. M., & Apostolakis, G. (1988). A critique of current practice for the use of expert opinions in probabilistic risk assessment. *Reliability Engineering & System Safety* 20.1, S. 63-85.
- Mudd, K., & Govern, J. M. (2004). Conformity to Misinformation and Time Delay Negatively Affect Eyewitness Confidence and Accuracy. *North American Journal of Psychology*, S. 227-238.
- Mullen, N., Charlton, J., Devlin, A., & Bedard, M. (2011). Simulator Validity: Behaviors Observed on the Simulator and on the Road. In D. Fisher, *Handbook of Driving Simulation for Engineering, Medicine and Psychology* (S. 13.1-13.18). Boca Raton, US: CRC Press.
- Müller, M. (2004). Risiko und Risikomanagement im Luftverkehr. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen*(7), S. 559-565.
- Nelson, B. (2004). *Accelerated testing: statistical method, test plans, and data analysis*. Hoboken, New Jersey, US: John Wiley & Sons Inc.
- Nolte, J. (2014). FSBD-Info. (G. d. Flugsicherung, Hrsg.) *Der Flugleiter* (6), S. 18.
- Ollman, R. (1966). Fast guesses in choice reaction time. *Psychonomic Science* 6.4, S. 155-156.
- Pelz, P., & Hedrich, P. (2015). *Unsicherheitsklassifizierung anhand einer Unsicherheitskarte (interner Bericht)*. Darmstadt, DE: Institut für Fluidsystemtechnik der TU Darmstadt.
- Pickup, L., Wilson, J., Norris, B., Mitchell, L., & Morrisroe, G. (2005). The Integrated Workload Scale (IWS): A new self-report tool to assess railway signaller workload. *Applied Ergonomics*(36.6), S. 681-693.
- Platz, K. (März 1991). Cooperative Air Traffic Management Concept (CATMAC). *EUROCONTROL Context*.
- Prevot, T., Lee, P., & Callantine, J. (2010). Human-in-the-loop evaluation of NextGen concepts in the Airspace Operations Laboratory. *Proceedings of the AIAA Modeling and Simulation Technologies (MST) Conference*. Toronto, CA.
- Purdy, G. (2010). ISO 31000:2009—Setting a New Standard for Risk Management. In *Risk Analysis* (Bd. 30.6, S. 881-886).
- Rasmussen, J. (1997). Risk Management in a Dynamic Behaviour, a Modelling Problem. *Safety Science*, 27, S. 183-214.
- Rastegary, H., & Landy, F. J. (1993). The Interactions Among Time Urgency, Uncertainty, and Time Pressure. In O. Svenson, & A. J. Maule, *Time Pressure and Stress in Human Judgment and Decision Making* (S. 217-235). New York, US: Plenum Press.
- Reason, J. (1990). *Human Error*. New York, US: Cambridge University Press.
-

- Sachs, L. (1999). *Angewandte Statistik: Anwendung statistischer Methoden, Auflage 9*. Springer.
- SAM-TF. (2004). *Air Navigation System Safety Assessment Methodology*. Brussels, BE: Eurocontrol.
- Sarris, V. (1990). *Methodologische Grundlagen der Experimentalpsychologie. 1. Erkenntnisgewinnung und Methodik der experimentellen Psychologie*. Reinhardt Verlag.
- Selye, H. (1976). Stress without Distress. In G. Serban, *Psychopathology of Human Adaptation* (S. 137-146). New York, US: Springer.
- SESAR. (2012). *European ATM Master Plan 2nd Edition*. Brüssel, BE: SESAR Joint Undertaking.
- SESAR. (Juni 2012). P16.1.1. *Accident Incident Modell V10-2*. Paris, FR.
- SESAR. (2013). *Validation / Verification of the Accident Incident Model (AIM)*. Brussels, BE: SESAR Joint Undertaking.
- SESAR. (2015). *Remotely Provided Air Traffic Service for Two Low Density Aerodromes Validation Report*. Brussels, BE: SESAR Joint Undertaking.
- Shapell, S., & Wiegmann, D. (1996). U.S. naval aviation mishaps 1977-92: Differences between single- and dual-piloted aircraft. *Aviation, Space, and Environmental Medicine*, S. 65-69.
- Shappell, S., Detwiler, C., Holcomb, K., Hackworth, C., Boquet, A., & Wiegmann, D. A. (2007). Human error and commercial aviation accidents: an analysis using the human factors analysis and classification system. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, S. 227-242.
- Shorrock, S. (2005). Who moved my (Swiss) cheese. *Aircraft and Aerospace*, S. 31-33.
- Shorrock, S. T. (2001). Assessing human error in air traffic management systems design: methodological issues. *Le travail humain*, S. 269-289.
- Stroeve, S. H., Blom, H., & Bakker, G. (2013). Contrasting safety assessments of a runway incursion scenario: Event sequence analysis versus multi-agent dynamic risk modelling. *Reliability Engineering & System Safety*(109), S. 133-149.
- Subotic, B., Ochieng, W., & Sträter, O. (2007). Recovery from equipment failures in ATC: Determination of the contextual factors. *Reliability Engineering and System Safety*(92.7), S. 858-870.
- Svenson, O., & Benson, L. (1993). On Experimental Instructions and the Inducement of Time Pressure Behaviour. In O. Svenson, & A. Maule, *Time Pressure and Stress in Human Judgement and Decision Making* (S. 157-166). New York, US: Plenum Press.
- Swain, A. D. (1990). Human reliability analysis: Need, status, trends and limitations. *Reliability Engineering & System Safety* (29.3), S. 301-313.
- Swain, A. D., & Guttman, H. E. (1983). *Handbook of human reliability analysis with emphasis on nuclear power plant applications*. NUREG.
- Swensson, R. G. (1972). The elusive tradeoff: Speed vs accuracy in visual discrimination tasks. *Perception & Psychophysics*(12.1), S. 16-32.
- The European Commission. (2010). *Commission regulation (EC) No 996/2010 - on the investigation and prevention of accidents and incidents in civil aviation and repealing Directive 94/56/EC*. Brussels, BE: Official Journal of the European Union.
- The European Commission. (2011). *Commission regulation (EC) No 1035/2011 - laying down common requirements for the provision of air navigation services*. *Official Journal of the European Union*, S. 23-41.
- Tversky, A., & Kahneman, D. (2000). Judgment under uncertainty: Heuristics and biases. In T. Connolly, *Judgment and decision making: An interdisciplinary reader* (S. 35-55).
- US Navy. (9. Mai 2016). *Current Mishap Definitions and Reporting Criteria*. Von http://www.public.navy.mil/navsafecen/Pages/statistics/mishap_def.aspx abgerufen
- Vogel, M., Thiel, C., & Fricke, H. (2012). Assessing the Air Traffic Control Safety Impact of Airline Pilot induced Latencies. *ATACCS'2012*. London, GB.

-
- Vu, K.-P. L., Minakata, K., & Nguyen, J. (2009). Situation awareness and performance of student versus experienced air traffic controllers. In G. Salvendy, & M. Smith, *Human Interface and the Management of Information. Information and Interaction* (S. 865-874). Springer Berlin Heidelberg.
- Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta psychologica*(41.1), S. 67-85.
- Wiegmann, D., & Shapell, S. (2001). *A Human Error Analysis of Commercial Aviation Accidents Using the Human Factors Analysis and Classification System*. Washington DC: Federal Aviation Administration.
- Windsor, H. (1928). Growing Wings for the Army. *Popular Mechanics*, 642-648.
- Yerkes, R., & Dodson, J. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, S. 459-482.

Danksagung

Bei der Ausarbeitung möchte ich mich bei meiner Familie bedanken, die die Zeit der Ausarbeitung mit viel Mithilfe und Geduld mitgetragen hat. Prof. Fricke hat meine Idee zur Zeitdruckstudie unterstützt und gefördert, wo immer er konnte. Für diese Freiheit zu forschen und die große fachliche Unterstützung möchte ich mich sehr bedanken. Mein Freund Christoph Thiel hat sich meine Ideen zur experimentellen Studie über Jahre hinweg angehört und mir mit viel Geduld und mit Hilfe seiner Erfahrung im Bereich Sicherheitsbewertung an entscheidenden Punkten aufgezeigt, die richtigen Fragen für die Untersuchung zu entwickeln. Ebenfalls möchte ich mich bei Prof. Pannasch bedanken, der mich mit hilfreichen Kommentaren versorgt hat. Die Labortreffen boten an seiner Professur die geeignete Plattform für Diskussionen mit Mitarbeitern des Fachbereiches Psychologie.

Anhang

Anhang A	Rollverkehrssimulator	I
Anhang B	Versuchsvorbereitungen	III
Anhang C	Kalibrierungsszenarien	IV
Anhang D	Kalibrierung des Leistungsmodells am Fluglotsenagenten	VI
Anhang E	Trainingsszenario	VIII
Anhang F	Versuchsplanung der Versuche im Januar 2014	IX
Anhang G	Szenariokonfigurationen	X
Anhang H	Analyse der Kalibrierungsdurchführung	XII
Anhang I	Analyse der Hauptversuche	XIV
Anhang J	Daten-CD	XXV

Anhang A Rollverkehrssimulator

Anhang A.1 Architektur

Der simulierte Betrieb wird mit Hilfe eines Rollverkehrssimulators realisiert, der das Verhalten von Luftfahrzeugen mikroskopisch abbildet und elementare Prozesse, wie z.B. Rollen, Starten, Landen und Pushback, auf Basis eines Luftfahrzeug-Verhaltensmodells berechnet. Die Initialisierung erfolgt anhand vordefinierter Flugpläne und eines Knoten-Kanten-Modells, in welchem Rollwege, Abflug- und Anflugpfade mit Hilfe von georeferenzierten Punkten zu einem Streckennetzwerk verbunden sind.

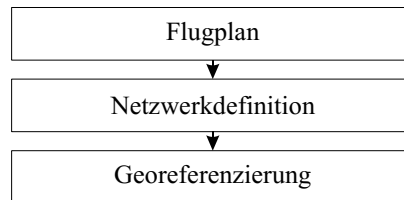


Abbildung: Flugplandefinition und Abhängigkeiten

Die automatisch generierten Flugpläne haben die Aufgabe, die generierte Verkehrslast über die Zeitachse steuern zu können. Dies erfolgt mit Hilfe von definierten Eingangskanälen, in dem pro Kanal Luftfahrzeuggruppen mit Verhaltensparametern zugeordnet sind. Diese Verteilungsparameter umfassen dabei den Zeitraum, die Anzahl und ein Parameter zur Varianz innerhalb des Zeitraums. Jeder Kanal ist einer Piste zugeordnet und enthält eine Kennung, ob es sich um eingehenden oder ausgehenden Verkehr handelt. Ebenso können Abhängigkeiten und Mindeststaffelungswerte im Anflug vorgegeben werden.

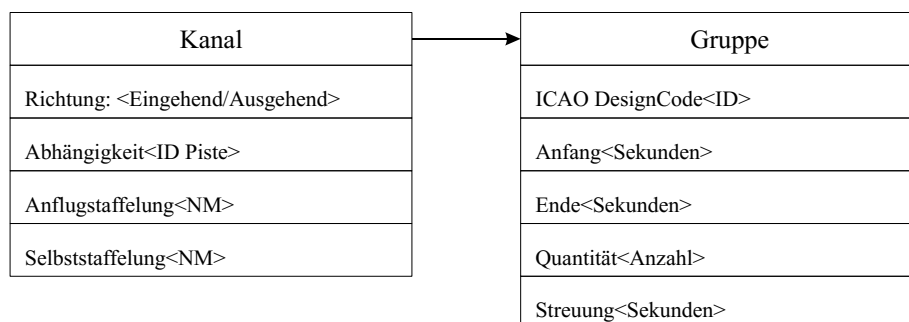


Abbildung: UML Diagramm der XML Flugplanstruktur

Der Eintrittszeitpunkt der Luftfahrzeuge einer Gruppe in die Simulation wird anhand einer Gleichverteilung innerhalb des Zeitintervalls bestimmt, der durch eine zufallsbasierte Varianz überlagert wird. Im Intervall T werden anhand einer solchen logischen Struktur n Flugbewegungen generiert, die eine Verkehrsdichte der Simulation wie folgt definiert:

$$\text{Verkehrsdichte} = \frac{n}{T}$$

Der Eintrittsort wird entsprechend der Piste und der Richtung festgelegt, wobei eingehender Verkehr auf den *Final Approach Fix* und ausgehender Verkehr auf eine Flugzeugstandfläche

bzw. Parkposition gesetzt werden. Sowohl die zu erreichende Flugzeugstandfläche des eingehenden Verkehrs, als auch der Aufrollpunkt des ausgehenden Verkehrs sind zufallsbasiert ermittelt und gleichverteilt innerhalb der zur Verfügung stehenden Menge an Möglichkeiten. Ebenso wird das Rufzeichen jedes Luftfahrzeugs zufallsbasiert generiert.

Nach Instanziierung der Verkehrsbewegungen aus den Flugplänen werden die jeweilig zugewiesenen Routen nach dem Prinzip der geringsten Kosten bestimmt und als Standardrouten gesetzt. Die Standardroute wird dem Probanden bei Erteilen der Rollfreigabe vorgeschlagen und kann bei Bedarf angepasst werden. Die Ermittlung der Standardroute erfolgt durch die Anwendung einer Kostenfunktion, in der eine Strecke im Netzwerk nach Anwenden des *Floyd Warshall Algorithmus* kostenoptimiert bestimmt wird. Diese Kosten sind i.d.R. die Länge eines Wegesegments oder werden innerhalb der Rollwege-Netzwerkdefinition dem jeweiligen Wegesegment manuell vordefiniert.

Das Verhalten der LFZ-Agenten ist in der Datei für die Luftfahrzeugparameter *AC_Defs.txt* anhand der folgenden Verhaltensparameter definiert:

- maximales Beschleunigen bei Roll-, Lande- und Startvorgang,
- maximales Abbremsen beim Roll- und Landevorgang,
- maximale Rollgeschwindigkeit,
- distanz zum Haltepunkt/Stopppunkt und
- mittlere Geschwindigkeit beim Aufsetzen beim Landevorgang.

Die Werte der Parameter werden nach Spezifikation von Nominal und Streuungswerten zufallsbasiert variiert.

Anhang B Versuchsvorbereitungen

Anhang B.1 Bewertung der Leistung des Kandidaten

Kandidatenbogen

Name: _____ Datum/Uhrzeit: _____

Durchführung: (Durchsatz & Zügigkeit, Sicherheit der Freigaben und der Bedienung)

Organisation: (Fachwissen bereits vorhanden? Pünktlichkeit?)

Fehlermeldungen der Simulation während der Durchführung: (Snapshot der Fehlermeldung in Textdatei oder Bilddatei? Notwendiger Neustart der Durchführung?)

Sonstige/ Notizen:

Anhang C Kalibrierungsszenarien

Anhang C.1 Baseline - Dauerlast

Die dazugehörige Versuchskonfiguration ist auf der Daten-CD, Anhang J, hinterlegt. Die Dateien in der Übersicht.

<i>Inhalt</i>	<i>Dateiname</i>
Konfigurationsdatei	SeparationLoad_Calibrate1.txt
Flugplandatei	FRA_Flightplan_SeparationLoad_Calibrate1.xml
Wegpunktdefinition	FRA_Navaids_Export.kml
Radarlayout	FRA_RadarLayout.kml
Hintergrundbild Surface Movement Manager	EDDF_black_0.92x_lm151213.svg
LFZ-Agentenverhalten	AC_Defs.txt
Streckennetzwerk	FRA_Routes.xml
Radar-Wegenetzwerk	FRA_RadarLayout.kml

Anhang C.2 Baseline - Spitzenlast

Die dazugehörige Versuchskonfiguration ist auf der Daten-CD, Anhang J, hinterlegt. Die Dateien in der Übersicht.

<i>Inhalt</i>	<i>Dateiname</i>
Konfigurationsdatei	SeparationLoad_Calibrate2.txt
Flugplandatei	FRA_Flightplan_SeparationLoad_Calibrate2.xml
Wegpunktdefinition	FRA_Navaids_Export.kml
Radarlayout	FRA_RadarLayout.kml
Hintergrundbild Surface Movement Manager	EDDF_black_0.92x_lm151213.svg
LFZ-Agentenverhalten	AC_Defs.txt
Streckennetzwerk	FRA_Routes.xml
Radar-Wegenetzwerk	FRA_RadarLayout.kml

Anhang C.3 ComPerf Kalibrierung

Die dazugehörige Versuchskonfiguration ist auf der Daten-CD, Anhang J, hinterlegt. Die Dateien in der Übersicht.

<i>Inhalt</i>	<i>Dateiname</i>
Konfigurationsdatei	ComPerfo_Calibrate.txt
Flugplandatei	FRA_Flightplan_ComPerfCalibrate.xml
Wegpunktdefinition	FRA_Navaids_Export.kml
Radarlayout	FRA_RadarLayout.kml
Hintergrundbild Surface Movement Manager	EDDF_black_0.92x_lm151213.svg

LFZ-Agentenverhalten	AC_Defs.txt
Streckennetzwerk	FRA_Routes.xml
Radar-Wegenetzwerk	FRA_RadarLayout.kml

Anhang D Kalibrierung des Leistungsmodells am Fluglotsenagenten

Anhang D.1 Verteilungsschätzung der Aktivzeit nach Maximum-Likelihood-Methode und Anpassungsgütetest nach Kolmogorow-Smirnow-Test

$t_L[s]$	Formparameter	Skalenparameter	P-Wert [%] ($n=1.000$)
0	1,92	592,5	99,3
10	1,76	581,0	68,1
20	1,75	612,62	7,0
30	1,81	639,43	28,6

Datei: weibull_fit_t_active.r

Anhang D.2 Abhängigkeit der mittleren Aktivzeit von der Latenzzeit und der Verkehrsmenge

$t_L[s]$ Sek./ LFZ	0	10	20	30
33,3	423,25	439,26	461,58	490,51
30,0	427,13	445,44	472,21	495,05
26,6	432,86	453,68	479,49	505,04
23,3	456,40	464,88	492,69	519,03
20,0	503,82	512,0	531,76	548,87

Datei: regression_TL2Taktiv_calibrate.r

Regressionsmodell: $\hat{t}_{aktiv} = a_0 + b_0 \cdot t_L + c_0 \cdot \frac{nk}{T}$

Ergebnis der linearen Regression:

Koeffizienten	Wert
a_0	312,3s
b_0	2,13
c_0	3441,3 s ² /LFZ

Anhang D.3 Abhängigkeit der Anzahl aktiver Verkehrsbewegungen von der Latenzzeit und der Verkehrsmenge

t_L [s] \n Sek./ LFZ	0	10	20	30
33,3	12,75	13,17	13,87	14,78
30,0	14,34	14,93	15,76	16,57
26,6	15,36	16,09	17,19	17,64
23,3	15,59	15,97	17,04	18,14
20,0	16,61	16,74	17,66	18,26

Datei: *regression_TL2Nsim_calibrate.r*

Regressionsmodell: $n_{simultan} = a_0 + b_0 \cdot t_L + c_0 \cdot \frac{n_k}{T}$

Ergebnis der linearen Regression:

Koeffizienten	Wert
a_0	8,5
b_0	$73,6e-3 \text{ s}^{-1}$
c_0	162,92 s/LFZ

Anhang E Trainingsszenario

Die dazugehörige Versuchskonfiguration ist auf der Daten CD, Anhang J, hinterlegt. Die Dateien in der Übersicht.

<i>Inhalt</i>	<i>Dateiname</i>
Konfigurationsdatei	Training_1.txt
Flugplandatei	FRA_Flightplan_Training.xml
Wegpunktdefinition	FRA_Navaids_Export.kml
Radarlayout	FRA_RadarLayout.kml
Hintergrundbild Surface Movement Manager	EDDF_black_0.92x_lm151213.svg
LFZ-Agentenverhalten	AC_Defs.txt
Streckennetzwerk	FRA_Routes.xml
Radar-Wegenetzwerk	FRA_RadarLayout.kml

Anhang F Versuchsplanung der Versuche im Januar 2014

KL – Kalibrierung

BL – Baseline

CP1 – ComPerf Zeitdrucklaststufe 1

CP2 – ComPerf Zeitdrucklaststufe 2

Versuchszeit		A	B	C
Datum	Uhrzeit			
10.1.14	9:03			BL
10.1.14	13:05	BL		
10.1.14	18:34		BL	
11.1.14	14:59	KL		
11.1.14	17:07		KL	
12.1.14	12:40			BL
14.1.14	14:34	CP2		
14.1.14	18:21		CP2	
17.1.14	14:16	CP1		
18.1.14	13:57	BL		
20.1.14	11:20			CP2
20.1.14	17:30	CP1		
21.1.14	9:33			CP1
21.1.14	15:54	CP2		
22.1.14	15:06			BL
22.1.14	18:22		CP1	
23.1.14	16:47			CP1
23.1.14	19:27		BL	
25.1.14	16:04		CP1	
27.1.14	12:32			KL
28.1.14	9:03			CP1
28.1.14	18:30		CP2	
30.1.14	11:10	CP2		
30.1.14	17:45			CP2

Anhang G Szenariokonfigurationen

Anhang G.1 Baseline

Die dazugehörige Versuchskonfiguration ist auf der Daten CD, Anhang J, hinterlegt. Die Dateien in der Übersicht.

<i>Inhalt</i>	<i>Dateiname</i>
Konfigurationsdatei	Baseline_1.txt
Flugplandatei	FRA_Flightplan_Baseline.xml
Wegpunktdefinition	FRA_Navaids_Export.kml
Radarlayout	FRA_RadarLayout.kml
Hintergrundbild Surface Movement Manager	EDDF_black_0.92x_lm151213.svg
LFZ-Agentenverhalten	AC_Defs.txt
Streckennetzwerk	FRA_Routes.xml
Radar-Wegenetzwerk	FRA_RadarLayout.kml

Anhang G.2 ComPerf 1

Die dazugehörige Versuchskonfiguration ist auf der Daten CD, Anhang J, hinterlegt. Die Dateien in der Übersicht.

<i>Inhalt</i>	<i>Dateiname</i>
Konfigurationsdatei	ComPerfo_1.txt
Flugplandatei	FRA_Flightplan_Baseline.xml
Wegpunktdefinition	FRA_Navaids_Export.kml
Radarlayout	FRA_RadarLayout.kml
Hintergrundbild Surface Movement Manager	EDDF_black_0.92x_lm151213.svg
LFZ-Agentenverhalten	AC_Defs.txt
Streckennetzwerk	FRA_Routes.xml
Radar-Wegenetzwerk	FRA_RadarLayout.kml

Anhang G.3 ComPerf 2

Die dazugehörige Versuchskonfiguration ist auf der Daten CD, Anhang J, hinterlegt. Die Dateien in der Übersicht.

<i>Inhalt</i>	<i>Dateiname</i>
Konfigurationsdatei	ComPerfo_2.txt
Flugplandatei	FRA_Flightplan_Baseline.xml
Wegpunktdefinition	FRA_Navaids_Export.kml
Radarlayout	FRA_RadarLayout.kml
Hintergrundbild Surface Movement Manager	EDDF_black_0.92x_lm151213.svg

LFZ-Agentenverhalten	AC_Defs.txt
Streckennetzwerk	FRA_Routes.xml
Radar-Wegenetzwerk	FRA_RadarLayout.kml

Anhang H Analyse der Kalibrierungsdurchführung

Anhang H.1 Latenzzeit zu Reaktionszeit

Datei: regression_tn_tL_calibration.r

<i>Proband</i>	<i>Bestimmtheitsmaß [%]</i>	<i>P-Wert [%]</i>
A	0,0	66,3
B	0,0	78,9
C	0,0	60,4

Anhang H.2 Latenzzeit zu mittleres Zeitbudget

Datei: regression_TL2TB_calibrate.r

<i>Proband</i>	<i>Bestimmtheitsmaß [%]</i>	<i>P-Wert [%]</i>
A	95,1	0,0***
B	92,2	0,0***
C	86,3	0,1***

Anhang H.3 Latenzzeit zur Häufigkeit der Zeitfehler

Datei: regression_TL2TE_calibrate.r

<i>Proband</i>	<i>Bestimmtheitsmaß [%]</i>	<i>P-Wert [%]</i>
A	99,4	0,0***
B	97,5	0,0***
C	37,9	10,4

Anhang H.4 Mittleres Zeitbudget zur Häufigkeit der Zeitfehler

Datei: regression_TB2TE_calibrate.r

<i>Proband</i>	<i>Bestimmtheitsmaß [%]</i>	<i>P-Wert [%]</i>
A	96,6	0,0***
B	92,7	0,0***
C	61,9	2,0*

Anhang H.5 Latenzzeit zu Arbeitsbeanspruchungsbefragung

Datei: regression_TL2WL_calibrate.r

<i>Proband</i>	<i>Bestimmtheitsmaß [%]</i>	<i>P-Wert [%]</i>
A	48,7	0,0***
B	43,3	0,0***
C	0,0	99,3

Anhang H.6 Latenzzeit zu Zeitdruckbefragung

Datei: regression_TL2TP_calibrate.r

<i>Proband</i>	<i>Bestimmtheitsmaß [%]</i>	<i>P-Wert [%]</i>
A	81,4	0,0***
B	90,7	0,0***
C	63,9	0,0***

Anhang H.7 Latenzzeit zu RI-Rate

Datei: regression_TL2RI_calibrate.r

<i>Proband</i>	<i>Bestimmtheitsmaß [%]</i>	<i>P-Wert [%]</i>
A	0,0	99,5
B	3,4	66,4
C	19,3	27,7

Anhang I Analyse der Hauptversuche

Anhang I.1 Vergleich der Aktivzeiten von Baseline und Lastenszenarien Probandenweise pro Runde

U-Test nach Mann-Whitney auf gleiche Tendenzen der Grundgesamtheit – Prüfung auf Gleichheit der Aktivzeiten-Verteilungen der Baseline gegenüber den Verteilungen unter Laststufen-bedingungen

Nullhypothese: $H_0: \tilde{t}_{aktiv\ CP} \geq \tilde{t}_{aktiv\ BL}$ (einseitig)

Datei: u-rank-test_T_aktiv_main.r

<i>Kandidat</i>	<i>Runde</i>	<i>Baseline</i>	<i>Lastenszenario</i>	<i>Laststufe</i>	<i>P-Wert U-Test [%]</i>
<i>A</i>	<i>1</i>	<i>10.1.14</i>	<i>17.1.14</i>	<i>1</i>	0,07***
<i>A</i>	<i>1</i>	<i>10.1.14</i>	<i>14.1.14</i>	<i>2</i>	0,0***
<i>B</i>	<i>1</i>	<i>10.1.14</i>	<i>22.1.14</i>	<i>1</i>	29,7
<i>B</i>	<i>1</i>	<i>10.1.14</i>	<i>14.1.14</i>	<i>2</i>	41,9
<i>C</i>	<i>1</i>	<i>12.1.14</i>	<i>21.1.14</i>	<i>1</i>	90,2
<i>C</i>	<i>1</i>	<i>12.1.14</i>	<i>20.1.14</i>	<i>2</i>	38,1
<i>A</i>	<i>2</i>	<i>18.1.14</i>	<i>20.1.14</i>	<i>1</i>	25,2
<i>A</i>	<i>2</i>	<i>18.1.14</i>	<i>30.1.14</i>	<i>2</i>	27,6
<i>B</i>	<i>2</i>	<i>23.1.14</i>	<i>25.1.14</i>	<i>1</i>	84,7
<i>B</i>	<i>2</i>	<i>23.1.14</i>	<i>28.1.14</i>	<i>2</i>	78,0
<i>C</i>	<i>2</i>	<i>22.1.14</i>	<i>28.1.14</i>	<i>1</i>	30,1
<i>C</i>	<i>2</i>	<i>22.1.14</i>	<i>30.1.14</i>	<i>2</i>	78,5

Anhang I.2 Abkürzungen und Einheiten der Messgrößen

Messgröße	Beschreibung	Entspricht unter Verwendung der Formelzeichen
DL	Nr. Runde [1,2]	
nZF	Häufigkeit Zeitfehler	ZF
nRI	Häufigkeit Runway Incursion gesamt	
nTL	Häufigkeit Runway Incursion – Taxi Landing	
nTS	Häufigkeit Runway Incursion – Taxi Starting	
nLS	Häufigkeit Runway Incursion – Landing Starting	
nSS	Häufigkeit Runway Incursion – Starting Starting	
nFr	Häufigkeit Freigaben	
mRT	Mittlere Reaktionszeit t_r [s]	$E(t_r)$
mAT	Mittlere Aktionszeit [s]	$E(t_{ak})$
nFG	Häufigkeit angeforderter Freigaben	
nRo	Häufigkeit erfolgreich modifizierter Rollwege	n_{rm}
mWL	Mittlere Arbeitsbeanspruchung nach IWS [0-7]	
saWL	Standardabweichung Arbeitsbeanspruchung nach IWS [0-7]	
mTP	Mittlere Zeitdruck nach IWS [0-7]	
saTP	Standardabweichung Zeitdruck nach IWS [0-7]	
nAC	Häufigkeit kontrollierter Luftfahrzeuge	
nSAC	Häufigkeit startender Luftfahrzeuge	
RIrate	Relative Häufigkeit von Runway Incursion pro Startfreigabe	RI-Rate
t_L	Latenzzeit des Fluglotsenagenten [s]	t_L
mTB	Mittleres angepasstes Zeitbudget [%]	\overline{TB}

Anhang I.3 Auswertungsübersicht

Es gelten die Abkürzungen, definiert in Anhang I.2.

Konf	Datum	Uhrzeit	Kandidat	DL	nZF	nRI	nTL	nTS	nLS	nSS	nFr	mRT	mAT	nFG	nRo	mWL	saWL	mTP	saTP	nAC	nSAC
BL	10.1.14	9:03	C	1		20	1	0	13	5	502	46,82	23,18	512	13					160	80
BL	10.1.14	13:05	A	1		3	2	0	0	1	519	58,49	23,54	528	24					160	80
BL	10.1.14	18:34	B	1		6	0	1	3	1	500	44,83	23,18	516	15					160	80
KL	11.1.14	14:59	A		20	9	2	1	5	1	603	35,91	23,22	619	22					205	103
KL	11.1.14	17:07	B		52	13	1	1	3	8	748	37,45	22,80	752	8					223	111
BL	12.1.14	12:40	C	1		3	0	0	1	2	507	45,55	23,94	522	17					160	80
CP2	14.1.14	14:34	A	1	12	4	2	0	2	0	473	36,63	21,25	483	19	2,79	0,67	3,86	0,74	158	79
CP2	14.1.14	18:21	B	1	14	17	3	1	8	5	609	35,21	21,69	626	17	3,56	1,12	3,89	1,70	188	94
CP1	17.1.14	14:16	A	1	8	4	0	1	3	0	557	38,99	19,92	576	22	2,88	0,83	3,12	1,53	176	88
BL	18.1.14	13:57	A	2		0	0	0	0	0	511	40,70	21,32	527	25					160	80
CP2	20.1.14	11:20	C	1	31	2	0	1	1	0	712	39,60	21,19	732	20	2,58	1,15	3,25	1,51	222	111
CP1	20.1.14	17:30	A	2	10	2	0	0	2	0	566	39,16	21,99	582	26	3,20	0,75	3,95	0,67	180	90
CP1	21.1.14	9:33	C	1	28	3	1	1	1	0	537	50,62	23,01	544	9	2,60	0,69	2,88	1,51	175	87
CP2	21.1.14	15:54	A	2	33	4	2	0	2	0	725	39,57	20,96	749	29	2,76	0,95	4,36	0,48	226	113
BL	22.1.14	15:06	C	2		6	3	0	2	1	503	47,20	22,05	515	17					160	80
CP1	22.1.14	18:22	B	1	10	3	0	2	1	0	573	37,59	23,47	590	24	3,59	0,84	3,24	1,99	180	90
CP1	23.1.14	16:47	C	2	15	2	2	0	0	0	594	42,28	21,69	610	17	3,14	0,99	3,33	1,25	190	95
BL	23.1.14	19:27	B	2		2	1	1	0	0	487	38,03	21,96	503	18					160	80
CP1	25.1.14	16:04	B	2	11	6	2	0	3	1	544	40,05	23,03	564	30	3,14	1,42	3,90	1,27	182	91
KL	27.1.14	12:32	C		37	3	3	0	0	0	684	40,22	22,84	699	20					215	107
CP1	28.1.14	9:03	C	2	15	0	0	0	0	0	617	39,97	20,87	635	16	2,71	0,88	3,24	1,44	190	95
CP2	28.1.14	18:30	B	2	6	6	1	0	4	1	535	35,30	22,25	544	26	3,21	1,00	2,74	1,52	172	86
CP2	30.1.14	11:10	A	2	13	4	1	1	2	0	592	36,26	22,06	604	18	3,16	0,81	3,68	1,17	186	93
CP2	30.1.14	17:45	C	2	22	2	1	0	1	0	646	49,14	22,63	667	14	3,05	0,88	3,91	2,48	204	102

Anhang I.5 Befragungsergebnisse der Arbeitsbeanspruchung - U-Test nach Mann-Whitney und T-Test

Datei: u-rank-test_WL_main.r

Alternativhypothese: $H_1: \tilde{x}_{CP1} \neq \tilde{x}_{CP2}$ (zweiseitig)

Intervall: 6 Minuten

Skala: 8-stufige „Integrated Workload Scale“ (IWS)

Proband	ComPerf 1	ComPerf 2	Runde	P- Wert [%]	
				U-Test	T-Test
A	17.1.2014	14.1.2014	1	65,2	73,2
B	22.1.2014	14.1.2014	1	75,6	92,4
C	21.1.2014	20.1.2014	1	95,7	95,3
A	20.1.2014	30.1.2014	2	91,6	87,1
B	25.1.2014	28.1.2014	2	95,5	86,6
C	28.1.2014	30.1.2014	2	23,4	23,5

Anhang I.6 Befragungsergebnisse des Zeitdrucks der Hauptversuche - U-Test nach Mann-Whitney und T-Test auf gleiche Tendenzen der Grundgesamtheit

Datei: u-rank-test_TP_main.r

Alternativhypothese: $H_1: \tilde{x}_{CP1} < \tilde{x}_{CP2}$ (einseitig)

Intervall: 6 Minuten

Skala: 8-stufige „Integrated Workload Scale“ (IWS)

Proband	ComPerf 1	ComPerf 2	Runde	P- Wert [%]	
				U-Test	T-Test
A	17.1.2014	14.1.2014	1	11,4	5,1
B	22.1.2014	14.1.2014	1	16,5	15,9
C	21.1.2014	20.1.2014	1	22,9	20,2
A	20.1.2014	30.1.2014	2	79,11	79,6
B	25.1.2014	28.1.2014	2	99,3	99,3
C	28.1.2014	30.1.2014	2	9,2	14,8

Anhang I.7 Vergleich der Reaktionszeiten der Lastenszenarien Probandenweise und pro Runde

Test1: U-Test nach Mann-Whitney auf gleiche Tendenzen der Grundgesamtheit – Prüfung auf Gleichheit der Reaktionszeit-Verteilungen der Baseline gegenüber den Verteilungen unter Laststufen-bedingungen

Alternativhypothese: $H_0: \tilde{t}_{r CP2} \geq \tilde{t}_{r CP1}$ (einseitig)

Datei: u-rank-test_T_r_main_laststufen.r

Kandidat	Runde	Lastenszenario		P-Wert [%]
		1	2	
A	1	17.1.14	14.1.14	51,1
B	1	22.1.14	14.1.14	43,7
C	1	21.1.14	20.1.14	9,4
A	2	20.1.14	30.1.14	6,2

<i>B</i>	2	25.1.14	28.1.14	0,996
<i>C</i>	2	28.1.14	30.1.14	91,6

Anhang I.8 Vergleich der Aktivzeiten der Lastenszenarien Probandenweise und pro Runde

Test1: U-Test nach Mann-Whitney auf gleiche Tendenzen der Grundgesamtheit – Prüfung auf Gleichheit der Aktivzeiten-Verteilungen unter variierenden Laststufenbedingungen

Test2: T-Test auf Mittelwertverschiebung – Prüfung auf Gleichheit der Aktivzeiten-Verteilungen der Baseline gegenüber den Verteilungen unter Laststufen-bedingungen

Alternativhypothese: $H_0: \tilde{t}_{aktiv\ CP2} \geq \tilde{t}_{aktiv\ CP1}$ (einseitig)

Datei: u-rank-test_T_aktiv_main_laststufen.r

<i>Kandidat</i>	<i>Runde</i>	<i>Lastenszenario</i>		<i>P-Wert</i>	
		1	2	<i>U-Test [%]</i>	<i>T-Test [%]</i>
<i>A</i>	1	17.1.14	14.1.14	1,2*	2,4*
<i>B</i>	1	22.1.14	14.1.14	61,1	67,2
<i>C</i>	1	21.1.14	20.1.14	3,9*	1,9*
<i>A</i>	2	20.1.14	30.1.14	57,7	55,7
<i>B</i>	2	25.1.14	28.1.14	45,3	42,4
<i>C</i>	2	28.1.14	30.1.14	92,7	93,2

Anhang I.9 Vergleich der Aktivzeiten von Baseline und Lastenszenarien auf Rundeübergreifende Trainingseffekte

Test 1: U-Test nach Mann-Whitney auf gleiche Tendenzen der Grundgesamtheit – Prüfung auf Gleichheit der Aktivzeiten-Verteilungen der Baseline gegenüber den Verteilungen unter Laststufen-bedingungen

Test 2: T-Test auf Mittelwertverschiebung – Prüfung auf Gleichheit der Aktivzeiten-Verteilungen der Baseline gegenüber den Verteilungen unter Laststufen-bedingungen

Alternativhypothese: $H_1: \tilde{t}_{aktiv\ R2} < \tilde{t}_{aktiv\ R1}$ (einseitig)

Datei: u-rank-test_T_aktiv_learning_main.r

<i>Kandidat</i>	<i>Szenario</i>	<i>Runde 1</i>	<i>Runde 2</i>	<i>P-Wert</i>	
				<i>U-Test [%]</i>	<i>T-Test [%]</i>
<i>A</i>	<i>BL</i>	10.1.14	18.1.14	0,8**	0,2**
<i>A</i>	<i>CP1</i>	17.1.14	20.1.14	20,3	27,9
<i>A</i>	<i>CP2</i>	14.1.14	30.1.14	94,1	94,6
<i>B</i>	<i>BL</i>	10.1.14	23.1.14	5,8	7,5
<i>B</i>	<i>CP1</i>	22.1.14	25.1.14	34,6	35,5
<i>B</i>	<i>CP2</i>	14.1.14	28.1.14	22,7	15,6
<i>C</i>	<i>BL</i>	12.1.14	22.1.14	45,5	51,1

<i>C</i>	<i>CP1</i>	21.1.14	28.1.14	2,3*	0,9**
<i>C</i>	<i>CP2</i>	20.1.14	30.1.14	88,1	87,3

Anhang I.10 Müdigkeitseffekte

Test: Lineare Regression auf Zeitabhängigkeit der Reaktionszeit

Datei: regression_t_r_to_time_main.r

<i>Kandidat</i>	<i>Runde</i>	<i>Lastenszenario</i>	<i>Laststufe</i>	<i>Bestimmtheitsmaß R² [%] (P-Wert [%])</i>
<i>A</i>	<i>1</i>	17.1.14	<i>1</i>	0,5 (8,5)
<i>A</i>	<i>1</i>	14.1.14	<i>2</i>	0,0 (86,6)
<i>B</i>	<i>1</i>	22.1.14	<i>1</i>	0,0 (81,2)
<i>B</i>	<i>1</i>	14.1.14	<i>2</i>	0,0 (89,0)
<i>C</i>	<i>1</i>	21.1.14	<i>1</i>	0,4 (12,1)
<i>C</i>	<i>1</i>	20.1.14	<i>2</i>	0,1 (34,8)
<i>A</i>	<i>2</i>	20.1.14	<i>1</i>	0,16 (33,5)
<i>A</i>	<i>2</i>	30.1.14	<i>2</i>	0,0 (33,4)
<i>B</i>	<i>2</i>	25.1.14	<i>1</i>	0,0 (64,6)
<i>B</i>	<i>2</i>	28.1.14	<i>2</i>	0,0 (99,8)
<i>C</i>	<i>2</i>	28.1.14	<i>1</i>	0,0 (70,2)
<i>C</i>	<i>2</i>	30.1.14	<i>2</i>	0,4 (11,7)

Anhang I.11 Analyse der Abhängigkeit der Häufigkeit der Aktivzeit und Zeitfehler von der Qualität und Reaktionszeit

Die Feststellung der Effektivität des Leistungsvergleiches anhand der Leistung wird innerhalb der Ergebnisse der Hauptversuche geprüft. So sollte das Wissen über die effektive Wirkung dieses Leistungsvergleiches eine gültige Randbedingung für alle Analysen der Daten bieten. Die Wirkungskette wird in diesem Rahmen von der Qualität der Flugplatzkontrolle bis zum Leistungsvergleich der Zeitdruck-Induktion auf Validität auf Gültigkeit. Die Stressreaktion ist dabei gemäß der implementierten Leistungsmessung und –vergleich der Zeitdruckinduktion anhand der folgenden Größe beobachtbar:

Mittlere Aktivzeit \hat{t}_{aktiv} als Resultierende des Qualitätsmerkmals $Q_{P\&S}$ und der Reaktionszeit t_r . Gemäß der Erläuterungen zum Qualitätsmerkmal $Q_{P\&S}$ des menschlichen Operateurs (Tabelle 8) wird als Maß der Qualität $Q_{P\&S}$ die Anzahl erfolgreicher Rollwegmodifikationen n_{rm} verwendet.

Weiterhin steht die Anzahl aktiver Verkehrsbewegungen (Abbildung 34), als Folge der mittleren Aktivzeit, gemäß dem implementierten Bestrafungsmechanismus in einem direkten Zusammenhang mit der Häufigkeit von Zeitfehlern.

Diese Zusammenhänge werden im Folgenden anhand der Messwerte der ComPerf-Konfiguration nachvollzogen, um den Zusammenhang zwischen den Leistungsebenen anhand

einer einfachen linearen Regression nachzuvollziehen. Dabei wurde nach den Laststufen unterschieden, um gleiche Randbedingungen für die Analyse herzustellen.

Die lineare Regressionsanalyse zeigt signifikante lineare Zusammenhänge entlang der Wirkungskette für Laststufen 1. Somit kann dem Kausalzusammenhang entlang der Wirkungskette des Leistungsmodells und Leistungsvergleich interne Validität unterstellt werden.

Die untere Abbildung zeigt anhand der Bestimmtheitsmaße der Laststufe 2 eine geringe Auswirkung der Routenmodifikationen auf die mittlere Aktivzeit zu haben. Der höhere Zeitdruck der Laststufe 2 priorisiert die Wichtigkeit dieses Mittels zur Optimierung geringer, bei der die Zeit zur Anweisung der Rollroute wenig Zeitersparnis gegenüber dem Standardrollweg bietet. Es kann ein Effekt verminderter Arbeitsqualität darstellen, bei der die Reaktionszeiten stärker ins Gewicht fallen. Die Zeitfehler stehen in einer 12,69 %igen Wahrscheinlichkeit in keiner Relation zur mittleren Aktivzeit. An dieser Stelle können bereits Effekte verstärkt auftretender RI-Raten als Folge des Zeitdrucks einwirken, die die Zeitbilanzen zugunsten der mittleren Aktivzeit verbessern. Diese Effekte zunehmenden Zeitdrucks werden im Folgenden näher untersucht.

Bestimmtheitsmaß R^2 der linearen Regression mit P-Wert mittlerer Aktivzeit \hat{t}_{aktiv} zu (1) Reaktionszeit t_r , (2) zur Häufigkeit der Rollwegmodifikationen n_{RM} und (3) Zeitfehler.

		Bestimmtheitsmaß R^2 [%] (P-Wert [%])		
Regressor		\hat{t}_r	n_{RM}	\hat{t}_{aktiv}
Regressand		\hat{t}_{aktiv}	\hat{t}_{aktiv}	ZF
Last	1	84,83 (0,91)**	-84,85 (0,91)**	80,32 (1,56)*
	2	74,59 (2,66)*	-33,25 (23,09)	58,05 % (7,8 %)

Datei: regression_mRT_nRO_nZF_2_Tactiv_main.r

Ausgabe R:

-----mRT zu mTA - Laststufe 1-----

R-square

[1] 0.8483272

P-Value

[1] "0.00910433556600112"

-----mRT zu mTA - Laststufe 2-----

R-square

[1] 0.7459475

P-Value

[1] "0.0266071175877184"

-----nRo zu mTA - Laststufe 1-----

R-square

[1] 0.848454

P-Value

[1] "0.00908867839935329"

-----nRo zu mTA - Laststufe 2-----

R-square

[1] 0.3325434

P-Value

[1] "0.230884533395587"

-----nZF zu mTA - Laststufe 1-----

R-square

[1] 0.8031733

P-Value

[1] "0.0156026594244709"

-----nZF zu mTA - Laststufe 2-----

R-square

[1] 0.5804648

P-Value

[1] "0.0782993685740512"

Anhang I.12 Lineare Regression der Latenzzeit zu Zeitbudget

Datei: regression_mRT_2_TB_main.r

Ausgabe R:

-----Regression tL zu TB, Hauptversuche -----

R-square

[1] 0.6642046

P-Value

[1] "0.00123999802355433"

Anhang I.13 Korrelationsanalyse der RI-Rate gegen Zeitfehler

Datei: correlation_ZF_2_RI_main.r

Ausgabe R:

----- Korrelation RI-Rate und Zeitfehler nach Pearson -----

Korrelationskoeffizient

[1] -0.5778

P-Value

	x	y
x	NA	0.2297
y	0.2297	NA

Anhang I.14 Lineare Regression der Latenzzeit zu RI Rate und Zeitfehler

Datei: regression_tl_2_RI-Rate_main.r

Ausgabe R:

-----Latenzzeit zu RI-Rate 1. und 2. Runde-----

R-square

[1] 0.1112522

P-Value

[1] "0.243852387908462"

-----Latenzzeit zu Zeitfehlern 1. und 2. Runde -----

R-square

[1] 0.08384102

P-Value

[1] "0.31531815997924"

-----Latenzzeit zu RI-Rate 1. Runde -----

R-square

[1] 0.171578

P-Value

[1] "0.414205841990243"

-----Latenzzeit zu Zeitfehlern 1. Runde-----

R-square

[1] 0.04194107

P-Value

[1] "0.69710186327695"

-----Latenzzeit zu RI-Rate 2. Runde -----

```

R-square
[1] 0.1059336
P-Value
[1] "0.431464158236103"
-----Latenzzeit zu Zeitfehlern 2. Runde-----
R-square
[1] 0.1333501
P-Value
[1] "0.373738136205043"

```

Anhang I.15 Lineare Regression und Korrelation der Latenzzeit zu RI Rate und Zeit nach Normierung

Datei: regression_correlation_Tl_2_P_Kehrwert_main.r

Ausgabe R:

```

-----Zusammenhang Latenzzeit zu RI-Rate vs ZF als Leistungsbemessung -----
[1] "lineare Regression des Kehrwert"
R-square
[1] 0.7012123
P-Value
[1] "0.0375157385623608"
[1] "Korrelation Kehrwert"
  x  y
x 1.00 -0.84
y -0.84 1.00

n= 6

```

```

P
  x  y
x  0.0375
y 0.0375
-----Zusammenhang Latenzzeit zu RI-Rate vs ZF als Leistungsbemessung -----
[1] "lineare Regression der Komplementärwahrscheinlichkeit"
R-square
[1] 0.4054579
P-Value
[1] "0.173955192092564"
[1] "Korrelation Komplementärwahrscheinlichkeit"
  x  y
x 1.00 -0.64
y -0.64 1.00

n= 6

```

```

P
  x  y
x  0.174
y 0.174

```

Anhang I.16 Bestimmung der Effektstärke nach Cohen

Datei: parameters_binomial.xlsx

Die Anzahl der Startfreigaben pro Versuchsdurchführung wurde zu jeder Zeitdrucklastbedingung gemittelt und die Anzahl der RI normiert. Die folgenden Formeln wurden zur Berechnung angewandt.

$$d = \frac{X_{n20} - X_{n30}}{\sigma_{Pooled}}$$

$$\sigma_{Pooled} = \sqrt{\frac{(n_1 - 1) \cdot \sigma_1^2 + (n_2 - 1) \cdot \sigma_2^2}{n_1 + n_2 - 2}}$$

Anhang J Daten-CD

Die beiliegende Daten-CD enthält das digitale Begleitmaterial, das zur Erstellung dieser Arbeit verwendet wurden. Sie sind entsprechend ihrer Verwendung nach Kapiteln und Abschnitten sortiert:

- Bilddateien
 - i. Corel-Draw v15, mit der Endung „cdr“
 - ii. Gnuplot-Skripte, mit der Endung „plt“ (Version 4.6)

Die Bilddateien sind in der Verzeichnisstruktur der Daten-CD unter dem jeweiligen Kapitel abgelegt. Die Corel-Bilddateien sind dabei in einem Unterverzeichnis mit dem Namen „Abb“ und der Abbildungsnummer abgelegt. Die Gnuplot-Bilder sind zumeist in dem Unterverzeichnis des jeweiligen Abschnitts, da die Ursprungsdaten auch zur Verwendung mit den Berechnungs-Skripten verwendet werden.

- Berechnungs-Skripte, für die Ausführung in dem Programm R (Version 3.1.1).
- Aufzeichnungsdaten des Versuches, entsprechend Abschnitt 3.2.6, mit der Endung „csv“. Der Dateiname enthält das Datum der Aufzeichnung und die dazugehörige Uhrzeit und das Kürzel des Probanden
 - i. Proband A: „NG“
 - ii. Proband B: „PM“
 - iii. Proband C: „PB“
- Software zur Auswertung und Aufbereitung der Daten, geschrieben in der Programmiersprache Java (Version 7) und entwickelt mit Hilfe der Entwicklungsumgebung Eclipse (Edition Juno)
- Szenariodateien, die für die Ausführung des Surface Movement Managers verwendet werden. Die Software des Surface Movement Managers selbst ist auf der Versionskontrolle „SVN“ der Professur Technologie und Logistik gespeichert:
<https://svnserve.vkw.tu-dresden.de/lme-t2l/svn/sman/>

Verzeichnis

Tag/ComPerfo_Januar_2014

Die Revision 615 entspricht der Version, wie sie für die Versuche verwendet worden ist.