Editorial

Liebe Leserinnen und Leser,

vor Ihnen liegt nunmehr die bereits vierzehnte Ausgabe des E-Journals **Anwendungen und Konzepte in der Wirtschaftsinformatik (AKWI)**.

Wir als Herausgeber möchten Sie nochmal darauf hinweisen, dass die älteren Ausgaben dieses Journals nach dem technischen Upgrade der Journal-Plattform immer noch im Archiv des Journals in kumulierter Form auffindbar sind.

Teil dieser Ausgabe sind wieder Zweitveröffentlichungen von Artikeln aus dem Bereich der Modellierung und Simulation, die von den Herausgebern als besonders interessant für die Leser dieses Journals empfunden wurden. Diese Artikel liegen in englischer Sprache vor und stammen von der ECMS 2021. Beispielhaft seien etwa ein Artikel zu einem evolutionären Optimierungsalgorithmus erwähnt, der auch im Umfeld der Wirtschafts-informatik Anwendung findet und der eine verbesserte Konvergenz zum globalen Minimum ermöglicht. Ein weiteres Beispiel behandelt eine Untersuchung zu einem MTO-Planungsproblem für saisonale Produkte, wobei eine Gewinnmaximierung angestrebt wird.

Wie schon in der dreizehnten Ausgabe, nimmt auch in der aktuellen Ausgabe das Geschäftsprozessmanagement einen zentralen Stellenwert ein. Diesmal wird dies durch einen Artikel aus dem Bereich des RPA (im Zusammenhang mit Retail-Prozessen) sowie durch einen Artikel zur Nutzung der Camunda Workflow-Engine zur Prozess-automatisierung abgedeckt. Abgerundet wird das Thema des Geschäftsprozess-managements durch einen Artikel, der als Praxisbericht in diese Ausgabe aufgenommen wurde und in dem eine Prozessdigitalisierung im Rahmen einer ERP-Einführung bei einem Anlagenbauer dargestellt wird.

Im Bereich der Optimierung/ Entscheidungsunterstützung im Bereich des Supply Chain Managements behandelt ein Artikel logistische Fallstudien im Zusammenhang mit der akademischen Software „LogisticsLab", die zum Lösen einer Reihe typischer Frage-stellungen aus dem Bereich des Supply Chain Managements dient, etwa dem Vehicle Routing Problem oder dem Network Flow Problem.

Eine eher betriebswirtschaftliche Fragestellung (Gesamtkostenbetrachtung von Firmenflugzeugen) wird in einem weiteren Artikel quantitativ betrachtet; hier wird insbe-sondere der monetäre Gegenwert der durch den Einsatz des Firmenflugzeugs eingesparten Reisezeit in die Gesamtkostenbetrachtung einbezogen.

Des Weiteren finden Sie auch wieder eine Reihe von Abschlussarbeiten, wobei hier der Schwerpunkt auf Fragen des cloudbasierten BI/ Data-Warehouseing liegt. In einer weiteren Abschlussarbeit wird jedoch auch der Zusammenhang zum analytischen CRM hergestellt.

Über Ihr Interesse an der Zeitschrift freuen wir uns und wünschen Ihnen Freude bei der Lektüre.

Regensburg, Fulda, Luzern und Wildau, im Dezember 2021.

Frank Herrmann, Norbert Ketterer, Konrad Marfurt und Christian Müller

Christian Müller

Konrad Marfurt

Norbert Ketterer

Frank Herrmann

FIRMENFLUGZEUGE – MONETÄRER GEGENWERT DER EINGESPARTEN REISEZEIT WIRD HÄUFIG UNTERSCHÄTZT

Matthias Bieberbach
Department of Industrial Engineering
Tor Vergata University of Rome
I-00133 Rome, Italy
E-Mail: matthias.bieberbach@t-online.de

KEYWORDS

Geschäftsluftfahrt, Werksflugverkehr, Opportunitätskosten.

ABSTRACT

Firmenflugzeuge werden durch die Öffentlichkeit oftmals argwöhnisch betrachtet. Als Gründe hierfür können die zahlenmäßig sehr hoch erscheinenden Kosten, fehlendes Wissen der Öffentlichkeit über die betrieblichen und wirtschaftlichen Vorteile, sowie an die Presse gelangte, fragwürdige Nutzungen genannt werden. Betrachtet man jedoch nicht nur die reinen Betriebskosten, sondern auch Faktoren, welche eben diese Kosten für den Betreiber in der Gesamtbetrachtung verringern, so wird das enorme Wertschöpfungspotential eines Firmenflugzeugs noch deutlicher. Durch Reisezeit verlorene Arbeitszeit kann ab einem gewissen Grad der Qualifizierung der Reisenden monetär durch einen Stundensatz ausgedrückt werden, sogenannte Opportunitätskosten. Dazu soll im Folgenden der Einfluss der Zeitersparnis auf die Gesamtbetriebskosten eines Firmenflugzeugs dargestellt werden. Dies geschieht im Kontext meiner PhD-Thesis, welche sich den break-even-Flugstunden eines Firmenflugzeugs als internes Reisemittel im Vergleich zu externen, kommerziellen Flugleistungen für das identische Reisevolumen beschäftigt. Weitere Optionen der Kostenreduzierung, wie beispielsweise eine Vercharterung an externe Kunden, sollen dabei außen vor bleiben. Basierend auf einem kurzen Abriss über Geschichte und Natur der Geschäftsluftfahrt, bezieht sich die Betrachtung der genannten Kosten ausdrücklich auf einen Eigenerwerb eines entsprechenden Flugzeugs und den Betrieb im Werksflugverkehr.

GESCHICHTE DER GESCHÄFTSLUFTFAHRT

Nachdem in den 1950er Jahren das erste reine Geschäftsreiseflugzeug von der Firma Learjet in den USA entwickelt worden war (Glancey 2005), blieb es zunächst einer kleinen, elitären Kundschaft vorbehalten (Budd und Graham 2009). Der breite Aufstieg und die Ausbreitung der Geschäftsluftfahrt begannen in den späten 70er-Jahren ebenfalls in den USA. Treibende Gründe waren dabei vorrangig eine zunehmende Unzufriedenheit mit dem Service etablierter Airlines (Budd und Graham 2009), sowie zahlreiche Neuentwicklungen von Flugzeugen speziell für diese Branche (Pope 2010), oftmals unter Verwendung von Technologien aus dem militärischen Flugzeugbau (Witt 2009). Parallel zu der von Linienflugzeugen (Wensveen 2011) erhöhte sich auch die Reichweite der Geschäftsluftfahrzeuge über die Zeit zusehends (Gulfstream 2018). Das zu dieser Zeit noch junge Internet ermöglichte ab den 90er-Jahren weltweite Buchungsplattformen (Gulfstream 2018) und trug somit die Erfolgswelle dieser Industrie in alle Welt. Die folgende Tabelle zeigt die weltweite Verteilung strahlgetriebener Geschäftsreiseflugzeuge im Jahr 2019 (JetNet 2019):

Tabelle 1: Weltweite Flottenverteilung 2019

Region	Anteil an der weltweiten Flotte [%]
USA	67
Europa	13
Südamerika	6
Mexiko	4
Mittel- und Fernost	3
Kanada	3
Zentralamerika	2
Afrika	1
Australien	1

Klar ersichtlich wird dabei der noch immer führende Einfluss des US-Markts.

VORTEILE DER GESCHÄFTSLUFTFAHRT

Seit dieser Zeit wurden Geschäftsreiseflugzeuge auch mehr und mehr als „Management-Tool" anerkannt (Gans 2011). Neben komfortablen Kabineneinrichtungen (Harrison 2013) werden heute vor allem Aspekte wie flexible Abflugzeiten, spontane Umplanungsoptionen und Vertraulichkeit geschätzt (Gans 2001; Hegmann 2015). Hinzu kommt das teils

immense Potential an Zeitersparnis durch separate Abfertigung und Nutzung regionaler Flughäfen mit der daraus resultierenden Reduzierung der Gesamtreisezeit im Vergleich zur Nutzung von Linienverbindungen. Insbesondere für Führungskräfte hat der Business Jet dadurch den Sprung aus der „Luxus-Ecke" geschafft, hin zu einem Werkzeug, welches trotz aller Kosten auch monetär zum Firmenerfolg beitragen kann (Entis 2014; NBAA 2009). Dies gilt für Führungskräfte von kleinen und mittleren Unternehmen (KMU) (Zielonka 2018) gleichermaßen wie für Top-Manager großer Konzerne oder den Transport zeitkritischer Güter, sowie hochqualifizierter Spezialisten zum Einsatzort (Cockpit 2018; Noack 2008). Diese Entwicklung ist mittlerweile weltweit zu beobachten, wenn auch mit regional unterschiedlichem Grad der Ausprägung (Thomalla 2005).

FLUGZEUGKLASSEN UND -KATEGORIEN

In der Bezeichnung der Größe und Reichweite aktueller Geschäftsreiseflugzeuge gibt es keine einheitliche Einteilung. Es werden zumeist Gewichts- und Reichweitenkategorisierungen vermischt.

Very Light-Klasse

Diese noch sehr junge, jedoch äußerst innovative (Tremesberger 2008) Klasse an Flugzeugen stellt die Hauptkonkurrenz für kleine Turboprop-Flugzeuge derselben Größe dar (Erdmann 2005) und spielt ihre Vorteile durch die Nutzung sehr kurzer Pisten und die Möglichkeit des Betriebs mit nur einem Piloten aus (Aerokurier 2007; Biermann 2007; Budd und Graham 2009), während diese Maschinen im Reiseflugzeug jedoch schneller fliegen können als propellergetriebene Muster vergleichbarer Größe. Typische Vertreter sind die Muster Cirrus Vision, ONE Aviation Eclipse and Honda HondaJet (Clarke 2018; Cox 2008).

Light-Klasse

Diese Klasse stellt den größten Anteil der weltweiten Flotte dar (Kern 2008). Verbunden mit ansprechendem Kabinenkomfort können aufgrund kurzer Start- und Landestrecken zahlreiche Flugplätze genutzt werden (Erdmann 2005). Typische Vertreter sind die Muster Embraer Phenom 300, Citation Citation CJ 1-4, Hawker 400 and Learjet 40/45 (Jetcraft 2015).

Medium-Klasse

Vertreter dieser Klasse stellen zahlenmäßig den zweitgrößten Anteil dar (Kern 2008). Sie verbinden deutlich erweiterten Komfort mit Ausstattung für effiziente Büroarbeit (Jetcraft 2015) und bieten zudem in gewissem Umfang bereits interkontinentale Reichweite (Erdmann 2005). Typische Vertreter sind die Muster Learjet 60, Falcon 2000 und die Challenger 300 Serie (Jetcraft 2015; Erdmann 2005).

Long Range-Klasse

Vertreter dieser Klasse bieten Flugzeiten jenseits von zehn Stunden (Padfield 2012) gepaart mit zahlreichen Optionen bezüglich der Innenausstattung ihrer großen Kabinen (Hegamnn 2016; Jetcraft 2015). Dieses Segment der Interkontinentalflugzeuge beherrschen die drei wichtigsten Hersteller: Bombardier mit Varianten ihrer Modelle Challenger 600 und Global 5000, Dassault mit der 900-Serie und Gulfstream mit ihrer 350/450-Serie (Erdmann 2005; Tremesberger 2008).

Ultra Long Range-Klasse

Muster in diesem Segment verschieben immer wieder die Höchstmarken in Bezug auf Reichweite und Komfort (Tremesberger 2008). Schnelle Internetverbindungen, welche Videokonferenzen zu jedem Zeitpunkt erlauben (Bombardier 2018), separate Schlafräume und sogar Duschen (Harrison 2013) erlauben immer längere Reisezeiten bei geringer körperlicher Belastung und der Möglichkeit effizienter Arbeit bei schier unbegrenzter Reisezeit (Rimmer 2018). Typische Vertreter sind die Muster Global 7500, Gulfstream 650ER und Falcon 7X/8X (Rimmer 2018).

Bizliner

Dieser Begriff ist eine Mischung aus *Business Jet* und *Airliner*. Flugzeuge dieser Kategorie bieten wesentlich mehr Platz als die zuvor genannten Klassen (Mose 2014). Etablierte Hersteller wie Airbus, Boeing oder Embraer bieten Ausstattungsoptionen für alle ihre gängigen Muster (Tremesberger 2008) mit nahezu unbegrenzten Optionen an (Mose 2014; Tremesberger 2008). Diese Flugzeuge sind jedoch eher der Kategorie des persönlichen Luxus zuzurechnen und spielen daher in der Betrachtung der Opportunitätskosten nahezu keine Rolle.

BETRIEBSKOSTEN

In vielen veröffentlichten Betrachtungen der Betriebskosten eines Business Jets werden nur direkt erfassbare Kosten aufaddiert. Keine Beachtung finden dabei jedoch zumeist diejenigen Faktoren, welche die Gesamtkosten reduzieren und damit die Rentabilität als Firmenjet hin zu einer geringeren Jahresflugstundenzahl verschieben können. Dies soll jedoch in diesem Artikel geschehen, mit dem Ziel, das Potential eines eigenen Flugzeugs für Unternehmen genauer beurteilen zu können.

Fixkosten und variable Kosten

Fixkosten eines Firmenflugzeugs, sind diejenigen Kosten, welche auf jeden Fall anfallen, unabhängig vom Flugaufkommen (Tennent 2008; Cannon und Richey 2012). Zudem können Fixkosten schwer kurzfristig reduziert werden (Doganis 2002), da beispielsweise zuerst das Flugzeug verkauft werden müsste. Variable Kosten hingegen entstehen erst durch ein spezifisches Flugereignis (McKnight 2010) und sind nicht im Voraus für eine bestimmte Abrechnungsperiode kalkulierbar (Cannon und Richey 2012). Jedoch können sie unmittelbar und kurzfristig reduziert werden (Doganis 2002), beispielsweise durch eine Reduzierung des Flugaufkommens.

Die folgende Tabelle kategorisiert alle relevanten Kosten eines Firmenflugzeugs nach ihrer Art:

Tabelle 2: Kategorisierung der Kosten

Kostenpunkt	Fixkosten	Variable Kosten
Administration	X	
Versicherung	X	
Crew - Reisekosten		X
Crew - Gehalt	X	
Crew - Training	X	
Dispatch / Flugplanung		X
Treibstoff		X
Ground-Handling		X
Finanzierungskosten	X	
Abschreibung	X	
MRO - Administration	X	
MRO - Arbeit und Material		X
Software und Lizenzen	X	

Tatsächliche Werte sind schwer zu recherchieren, jedoch bieten diverse Brancheninformationen und Veröffentlichungen innerhalb der Geschäftsluftfahrt einzelne Zahlen zu den diversen Kostenpunkten, so dass sich mithilfe zahlreicher Quellen die folgenden Werte in USD als realistische Durchschnittswerte für die Vertreter der Light-Klasse ergeben, und zwar im privaten Werksflugverkehr ohne Managementvertrag mit Flugbetrieb in Europa (Aviationjobsearch.com 2019; Betteraviationjobs.com 2020; Bizjetjobs.com 2019; Collogan 2014; Fuelworx 2021; Globalair.com 2021; JetNet 2021; Newby 2015; Operations Planning Guide 2018; Raouna 2014; Sherpareport 2013; Sprouts.aero (2021); Thurber 2008; Wyndham 2020; Young-Brown 2015; Young-Brown 2016). Die Fixkosten fallen dabei pro Jahr an, die variablen Kosten pro Flugstunde:

Tabelle 3 – Betriebskosten Light-Klasse

Light-Klasse	Fixkosten pro Jahr [USD]	Variable Kosten pro Flugstunde [USD]
Administration	5.000	
Versicherung	25.000	
Crew - Reisekosten		300
Crew - Gehalt	132.000	
Crew - Training	30.000	
Dispatch / Flugplanung		60
Treibstoff		650
Ground-Handling		50
Finanzierungskosten	300.000	
Abschreibung	500.000	
MRO - Administration	95.000	
MRO – Arbeit und Material		900
Software und Lizenzen	5.000	
SUMME	1.092.000	1960

Basis für die recherchierten variablen Kosten sind dabei 200 Flugstunden pro Jahr, so dass bei höherem Flugaufkommen eine gewisse Abweichung erwartet werden kann. Diese steht jedoch der generellen Darstellung nicht entgegen. Aufgrund der Internationalität der verwendeten Quellen und des Handels von Treibstoff in USD, wurden alle Zahlen in USD angegeben. Reisekosten und Kosten für Parken und Hangarierung wurden basierend auf der Annahme ausgewählt, dass das Flugzeug hauptsächlich von der Heimatbasis aus operiert und regelmäßig dahin zurückkehrt. Für die angesetzten 15 Einsatztage im Monat wurde mit nur einer Crew kalkuliert, bei einer höheren Zahl von Arbeitstagen erhöhen sich entweder die Fixkosten durch zusätzliche Festangestellte oder die variablen Kosten durch den Einsatz von Freelancern. Als Finanzierungskosten wurde eine interne Finanzierung angenommen und ein entgangener Gewinn von 3% auf das nun gebundene Kapital in Form des Kaufpreises angesetzt. Für die Abschreibung wurde der Kaufpreis über branchenübliche 20 Jahre verteilt. In der genannten Rate für MRO (Maintenance, Repair and Overhaul) sind Reserven für größere Überholungen bereits einkalkuliert.

Opportunitätskosten

Wie eingangs erwähnt, lässt sich aufgrund vermeidbarer Reisezeit verlorene Arbeitszeit (Behrmann und Stepp 2001) durchaus in geldlichem Gegenwert ausdrücken,

als sogenannte Opportunitätskosten. Umgekehrt kann somit eingesparte Reisezeit zusammen mit unterwegs gewonnener, effektiver Arbeitszeit (Jorge-Calderón 2014) die Betriebskosten eines Firmenjets in der Gesamtbetrachtung reduzieren. Der europäische Dachverband EBAA hat Zahlen zur eingesparten Reisezeit veröffentlicht, gültig für die beiden kleinsten Klassen (EBAA 2018). Exemplarisch soll im Folgenden die Light-Klasse betrachtet werden. Die durchschnittliche Flugdauer pro Strecke wurde zudem mit 50% der veröffentlichten maximalen Reichweite von 3h (Ellis 2019) dieser Klasse abgeschätzt, die zusätzlich gewonnene Arbeitszeit unter Abzug typischer Zeiten für Ab- und Anflug. Für die betrachtete Flugzeugklasse und diese durchschnittliche Flugzeit pro Strecke führen zu einer jährlichen Anzahl an Flugbewegungen von $2/_3$ der Jahresflugstundenzahl. Der Gesamtgewinn an effektiver Arbeitszeit pro Flugstrecke ergibt sich danach wie folgt:

Tabelle 4 – Zeitersparnis pro Strecke am Beispiel Light-Klasse

Aspekt	Zeitersparnis [h]
Reisezeit	2
Gewonnene Arbeitszeit	1
Gesamter Gewinn an effektiver Arbeitszeit	**3**

Bei der Betrachtung der Kosten pro Flugstunden eines firmeneigenen Flugzeugs muss nach jährlicher Gesamtflugstundenzahl differenziert werden, da die Betriebskosten pro Flugstunde durch die anteilige Verteilung der Fixkosten mit steigender Jahresflugstundenzahl sinken. Das folgende Diagramm zeigt diesen Effekt qualitativ:
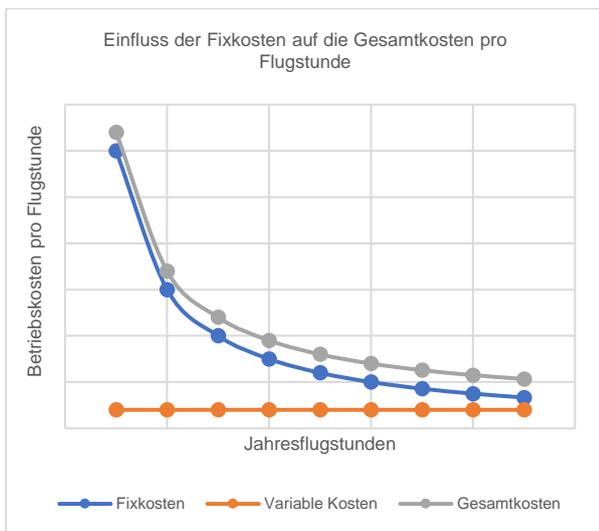

Diagramm 1 – Verteilung der Fixkosten pro Flugstunde

Die anteilige Ersparnis durch die Opportunitätskosten hängt folglich auch von der Zahl der Jahresflugstunden ab. Das folgende Diagramm zeigt das Einsparungspotential der Betriebskosten pro Flugstunde durch prozentualen Bezug der Opportunitätskosten auf die Gesamtkosten pro Flugstunde und unter Verwendung der oben dargestellten Kosten und verschiedener Jahresflugstundenzahlen. Die x-Achse zeigt dabei den Gesamtwert der Opportunitätskosten in USD, die y-Achse den prozentualen Anteil der Einsparung aufgrund entsprechender Opportunitätskosten. Ausgehend vom flachsten Graphen werden nach links hin die Daten für Jahresflugstunden von 100 bis 1000 in 100er-Schritten dargestellt:
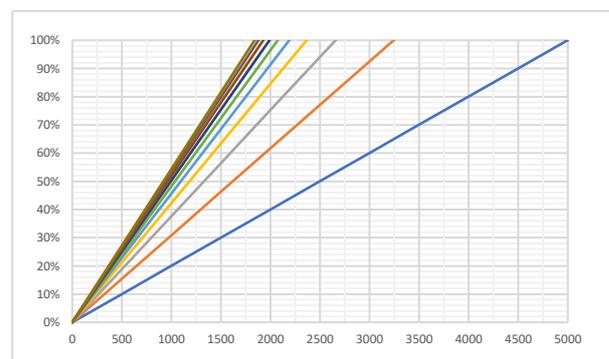

Diagramm 2 – Einfluss der Opportunitätskosten

Da wie zuvor erwähnt, die genutzten Werte für die variablen Kosten auf eine bestimmte Jahresflugstundenzahl bezogen sind, ist weit abseits dieser Jahresfluggleistung ein gewisser Fehler nicht zu vermeiden. Vom Trend her steht dies jedoch der Gesamtdarstellung nicht entgegen. Die verwendeten Werte für die Opportunitätskosten stehen dabei als gesamter Stundensatz aller Reisenden und es obliegt jedem Nutzer, sich selbst Gedanken zum monetären Wert der Arbeitszeit seiner eigenen Mitarbeiter zu machen. Zur besseren Vergleichbarkeit wurde der Stundensatz ebenfalls in USD angegeben. Es kann sich dabei um den fiktiven Stundensatz eines einzelnen Topmanagers handeln oder auch um das Produkt aus dem Stundensatz von schnell zu transportierenden Mitarbeitern und der entsprechenden Passagieranzahl pro Strecke. Folglich gilt: je mehr Passagiere pro Strecke transportiert werden, desto höher ist das Einsparpotential. Voraussetzung dafür ist jedoch, dass alle mitreisenden Personen auch im Falle der Nutzung eingekaufter Flugleistungen mitreisen würden, ihre Reisetätigkeit also nicht nur aufgrund freier Kapazitäten erfolgt.

CONCLUSION

Ausgedrückt in konkretem, finanziellem Vorteil kann also eingesparte Reisezeit, als einer der am stärksten propagierten Vorteile der Geschäftsluftfahrt, die tatsächlichen Kosten pro Flugstunde abhängig von Passagieranzahl und dem wirtschaftlichen Leistungspotential der Reisenden somit deutlich senken. Unter Umständen können die Betriebskosten eines firmeneigenen Flugzeugs somit sogar komplett kompensiert werden. Zahlreiche Beispiele aus der deutschen Industrie untermauern diese Vorteile, auch wenn Informationen über konkrete Zahlen und Betriebsstrukturen nur sehr spärlich nach außen gelangen. So unterhalten laut der Homepage der GBAA, dem Interessenverband der deutschen Geschäftsluftfahrt, unter anderem Volkswagen, Würth, Bosch und Bertelsmann eigene Flugzeuge und auch die Maschinen der Telekom, von BASF, BMW oder Lidl sind regelmäßig an deutschen und europäischen Flughäfen zu sehen. Von außen sind sie selten als solche Firmenvertreter zu erkennen, doch wenn man genauer hinsieht, so findet sich oftmals versteckt in der Registrierung ein kleiner Hinweis auf ihr Mutterunternehmen.

AUTHOR BIOGRAPHIES

Matthias Bieberbach wurde 1978 in Nürnberg geboren und begann seine berufliche Laufbahn 1997 als Offiziersanwärter in der Luftwaffe. Als junger Offizier absolvierte er das Studium der Luft- und Raumfahrttechnik an der Universität der Bundeswehr in München und wurde dann an der Verkehrsfliegerschule der Lufthansa in Bremen zum Verkehrsflugzeugführer ausgebildet. Als Pilot der Flugbereitschaft BMVg war er Mitglied des Regierungsfliegergeschwaders und sammelte Erfahrung auf Geschäftsreiseflugzeugen des Herstellers Bombardier. Parallel erlangte er 2010 den Abschluss als MBA Luftverkehrsmanagement an der TH Wildau. Seit dem Ende seiner militärischen Laufbahn ist er als Kapitän auf der Bombardier Global Serie in der zivilen Luftfahrt tätig. Mit dem Ziel, seine weitreichenden und in ihrer Kombination nahezu einzigartigen technischen, organisatorischen und operationellen Erfahrungen zu bündeln, arbeitet er derzeit an der Tor Vergata Universität Rom an seiner PhD-Thesis. Seine e-mail-Adresse lautet: matthias.bieberbach@t-online.de und sein berufliches Profil ist verfügbar unter: linkedin.com/in/mbieberbach.

REFERENCES

Aerokurier. 2007. „Single Pilot Operation: Die neue Leichtigkeit der Minijets". https://www.aerokurier.de/general-aviation/single-pilot-operation-die-neue-leichtigkeit-der-minijets/477411 [Zugriff 27/12/2018]

Aviationjobsearch.com. 2019. "Salary report 2019". https://blog.aviationjobsearch.com/wp-content/uploads/2020/02/AJS-Salary-report-2018-19-FINAL.pdf [Zugriff 26/11/2020]

Behrmann, D. und Stepp, M. 2001. "Business aviation in today's economy: A guide to the analysis of business aircraft use, benefits and effects on shareholder value". https://nbaa.org/wp-content/uploads/2018/02/AndersenPart02.pdf, [Zugriff 05/01/2021]

Betteraviationjobs.com. 2020. "Corporate jet pilot salary study (2020)". https://betteraviationjobs.com/news/corporate-jet-pilot-salary-study-2020/ [Zugriff 27/11/2020]

Biermann, T. 2007 „Wie fliegen wir morgen? – Neue Tendenzen im Luftverkehr". In TFH Wildau – Wissenschaftliche Beiträge 2007

Bizjetjobs.com. 2019. "2019 Pilot salary survey results 2019". https://bizjetjobs.com/pilot-salary-survey/#chief [Zugriff 24/11/2020]

Bombardier. 2018a. „Ka-band". https://businessaircraft.bombardier.com/sites/default/files/2018-10/GlobalKaBand.pdf, Zugriff 03/11/2018

Budd, L. und Graham, B. 2009. "Unintended trajectories: liberalization and the geographies of private business flight". In Journal of Transport Geography Vol. 17(4), p. 285-292

Budd, L. und Ison, S. 2017. Air transport management, Oxon

Cannon, J. und Richey, F. 2012. Practical applications in business aviation management, Plymouth

Clarke, C. 2018. "The very best light jets in the sky today". https://www.popularmechanics.com/flight/g21098242/best-light-jet-airplanes/ [Zugriff 28/12/2018]

Cockpit. 2018. „Die Business Aviation im Dienste der Wirtschaft". http://www.cockpit.aero/rubriken/detailseite/news/die-business-aviation-im-dienste-der-wirtschaft/?no_cache=1 [Zugriff 02/11/2018]

Collogan, D. 2014. "What drives jet fuel prices?". https://www.bizavadvisor.com/what-drives-jet-fuel-prices/ [Zugriff 05/01/2021]

Cox, B. 2008. "The day of the personal jet". https://www.planeandpilotmag.com/article/the-day-of-the-personal-jet/ [Zugriff 27/12/2018]

Doganis, R. 2002. Flying off course – The economics of international airlines, New York

EBAA. 2018b. "European Business Aviation – Economic value & business benefits". https://www.ebaa.org/app/uploads/2018/01/EBAA-Economic-report-2017_compressed.pdf [Zugriff 15/11/2018]

Ellis, C. 2019. "The cost of flight on a private jet in the USA". https://www.aircharterserviceusa.com/about-us/news-features/blog/how-much-does-it-cost-to-charter-a-private-jet [Zugriff 11/12/2020]

Entis, L. 2014. "When it makes sense to take a private jet". https://www.entrepreneur.com/article/233067 [Zugriff 02/11/2018]

Erdmann, C. u. a. 2005. Business Jets: Markt, Betreibermodelle, Eigentümer, Markttrends, Gebrauchtmarkt, HSH-Nordbank Branchenstudie, Hamburg

Fuelworx. 2021. Datenabfrage über privaten Zugang. https://prices.fuelworx.aero/login.php [Zugriff 12/01/2021]

Gans, B. 2001. „Raus aus der Glamour-Ecke". In AERO 9/2001, S. 38

Glancey, J. 2015. "The Learjet: The private jet that changed travel". http://www.bbc.com/culture/story/20150813-the-ultimate-in-personal-luxury [Zugriff 04/11/2018]

Gollan, D. 2015. "A beginners guide to choosing your private jet". https://www.forbes.com/sites/douggollan/2015/11/15/a-beginners-guide-to-choosing-your-private-jet/#7b4ccdf55076 [Zugriff 26/12/2018]

Gulfstream. 2018. "First transatlantic business jet flight". https://www.youtube.com/watch?v=gaJQJfowSjc [Zugriff 21/09/2019]

Harrison, K. 2011. "Outfitting your aircraft`s cabin". https://www.bjtonline.com/business-jet-news/outfitting-your-aircrafts-cabin [Zugriff 26/12/2018]

Hartmann, P. (2017). „Privatjet verchartern oder nicht?". https://blog.premiumjet.eu/de/privatjet-verchartern-oder-nicht/ [Zugriff 21/02/2021]

Hegmann, G. 2015. „Sixt gründet eine Fluggesellschaft für sich selbst". https://www.welt.de/wirtschaft/article147196021/Sixt-gruendet-eine-Fluggesellschaft-fuer-sich-selbst.html [Zugriff 21/01/2021]

Hegmann, G. 2016. „Die extravaganten Wellness-Jets der Superreichen". https://www.welt.de/wirtschaft/article155975401/Die-extravaganten-Wellness-Jets-der-Superreichen.html [Zugriff 26/12/2018]

Humaira, N. u. a. 2016. "Aircraft scheduled airframe maintenance and downtime integrated cost model". In Advances in Operations Research, February 2016

Jetcraft. 2015. "The basics of private jet categories". https://www.jetcraft.com/jetstream/2015/06/basics-private-jet-categories-2/ [Zugriff 28/12/2018]

JetNet. 2019. "*JetNet iQ market report – 2019*". https://www.jetnet.com/research/reports/jetnet-iq-market-report-2019.html [Zugriff 10/05/2021]

JetNet. 2021. Datenabfrage über privaten Zugang. https://www.jetnetevolution.com/Default.aspx [Zugriff 13/02/2021]

Jorge-Calderón, D. 2014. Aviation investment, Farnham

Kern, J.-H. 2008. Analyse der Wettbewerber auf dem Markt des kommerziellen Bedarfsflugverkehrs, München

Liberty Jet. 2020. "Jet performance and specifications". https://www.libertyjet.com/private-jet-fleet.aspx [Zugriff 08/12/2020]

McKnight, P. 2010. Airline economics. In Gleich, R. / Wald, A. (2010), Introduction to aviation management, Berlin

Mensen, H. 2003. Handbuch der Luftfahrt – 1. Auflage, Berlin

Mose, J. 2014. "Buying your first bizliner". In Business Aviation Advisor April 2015, S. 6ff

NBAA. 2009. "Explaining why your company relies on business aviation". https://nbaa.org/flight-department-administration/tax-issues/federal-taxes/explaining-why-your-company-relies-on-business-aviation-questions-and-answers/ [Zugriff 02/11/2018]

Newby, N. 2015. "The costs to own and operate a Cessna CJ3". https://www.sherpareport.com/aircraft/costs-cessna-cj3.html [Zugriff 08/12/2020]

Noack, H.-C. 2008. „Das fliegende Arbeitszimmer". In Frankfurter Allgemeine Zeitung 13/11/2008, S. 20

Operations Planning Guide. 2018. "How to use the operating costs guide". In Business & Commercial Aviation 08/2018, Vol. 114, S. 42ff

Padfield, R. 2012. "Jets for the long haul". https://www.barrons.com/articles/SB50001424053111903935304577380262461370738 [Zugriff 28/12/2018]

Picheta, R. 2018. "Boeing launches business jet capable of world`s longest flight". https://edition.cnn.com/travel/amp/boeing-business-jet-longest-flight-scli-intl/index.html [Zugriff 28/12/2018]

Pope, J. 2010. "Editor's Desk: What business aviation's history tells us about its future". In Business Jet Traveller 5/2010, S. 4

Raouna, K. 2014. "How to become a private flight attendant". https://www.careeraddict.com/become-a-vip-flight-attendant [Zugriff 07/02/2019]

Rimmer, D. 2018 „Bombardier launches new ultra-long-range jets". https://www.thebalancecareers.com/bombardier-launches-new-ultra-long-range-jets-282805 [Zugriff 29/12/2018]

Sherpareport. 2013. "The costs of buying and operating an Embraer Phenom 300". https://www.sherpareport.com/aircraft/aircraft-overview/costs-embraer-phenom-300.html [Zugriff 27/11/2020]

Sprouts.aero. 2021. Email-Konversation mit Herrn Schaller

Thomalla, V. 2005. „Bestandsaufnahme". In Flugrevue 2/2005, S. 38 – 41

Tennent, J. 2008. Guide to financial management - 1st edition, London

Tremesberger, D. 2008. „Der europäische Markt für Business Aviation - Analyse des Angebots im internationalen Vergleich unter Berücksichtigung von Einflussfaktoren auf das Wachstum". In Schriftenreihe des Instituts für Transportwirtschaft und Logistik, Wirtschaftsuniversität Wien, Nr. 1 (2008 SCM)

Wensveen, J. 2011. Air transportation – A management perspective, Farnham

Witt, D. 2009. "Fly-by-wire evolution and the Gulfstream G650". In Professional Pilot 5/2009, S. 77f

Wyndham, D. 2020. "What does it cost to operate a light jet". https://www.avbuyer.com/articles/operating-costs/what-does-it-cost-to-operate-a-light-jet-112746 [Zugriff 24/11/2020]

Young-Brown, F. 2015. "Fuel burn rates for private aircraft. https://www.sherpareport.com/aircraft/fuel-burn-private-aircraft.html [Zugriff 05/01/2021]

Young-Brown, F. 2016. "The costs to own and operate a Gulfstream G450". https://www.sherpareport.com/aircraft/costs-gulfstream-g450.html [Zugriff 27/11/2020]

Zielonka, G. 2018. „Trendwende in der Geschäftsluftfahrt". In Der Mobilitätsmanager Mai/Juni 2018, S. 16f

DIFFERENTIAL EVOLUTION ALGORITHM IN MODELS OF TECHNICAL OPTIMIZATION

Roman Knobloch and Jaroslav Mlýnek
Department of Mathematics
Technical University of Liberec
Studentská 2, 46117 Liberec
The Czech Republic
E-mail: roman.knobloch@tul.cz, jaroslav.mlynek@tul.cz

KEYWORDS

Optimization, search space, cost function, differential evolution algorithm, global convergence, asymptotic convergence, mathematical model.

ABSTRACT

At present, evolutionary optimization algorithms are increasingly used in the development of new technological processes. Evolutionary algorithms often allow the optimization procedure to be performed even in cases where classical optimization algorithms fail (e.g. gradient methods) and where an acceptable solution is sufficient to solve the optimization task. The article focuses on possibilities of using a differential evolution algorithm in the optimization process. This algorithm is often referred to in the literature as a global optimization procedure. However, we show by means of a practical example that the convergence of the classic differential algorithm to the global extreme is not generally assured and is largely dependent on the specific cost function. To remove this weakness, we designed a modified version of the differential evolution algorithm. The improved version, named the modified differential evolution algorithm, is described in the article. It is possible to prove asymptotic convergence to the global minimum of the cost function for the modified version of the algorithm.

INTRODUCTION

New technological procedures are often developed using mathematical models describing the essential features of the solved problem. The model is then used to transform the real world problem into an optimization task. Strong assumptions are often required when using classic optimization methods (e.g. convexity of the searched space, convexity of the evaluation function, knowledge of the appropriate position of the initial solution). Otherwise, these methods do not often lead to the required solution.

Evolutionary optimization algorithms are primarily utilized in situations when other usual methods fail to converge to the optimized state. Recently, use of the evolutionary optimization algorithms has been considerably expanding, see e.g. (Simon 2013), (Affenzeller et al. 2009)). The evolutionary algorithms are in particular appropriate for problems with a complicated structure of the search space and in case of intricate cost functions. Evolutionary algorithms are in general more computationally demanding and they are therefore suitable for calculations that are not time limited (e.g. off-line calculations of trajectories of an industrial robot, see (Mlýnek et al. 2020)). Their use in time critical calculations is rather limited. For example, their utilization for online decision making processes (e.g. online calculations of trajectories of industrial robot depending on the evaluation of current conditions) is not so frequent. Nevertheless, parallel programming tools are often used to speed up calculations with good results. Nowadays, parallel programming tools form a part of most used programming languages.

The differential evolution algorithm is one of the frequently used algorithms for solving practical optimization tasks. This algorithm was first introduced by Storn and Price in (Storn and Price 1997) and (Price et al. 2005). This algorithm is often referred to as a global optimization method (see (Storn and Price 1997), (Price 1996)). However, such statements are always justified. We demonstrate by an example of a specific cost function that this algorithm is prone to premature local convergence and its convergence to the minimum of the cost function is not assured. The issue of suitable choice of optional algorithm parameters is solved, for example, in (Červenka and Boudná 2018). In this article we propose a suitable modification of the differential evolution algorithm that eliminates the premature convergence to a local minimum. Additionally, it is possible to prove asymptotic convergence to the global minimum of the cost function.

The differential evolution algorithms now constitute a larger group of similar algorithms that differ in implementation details. We concentrate on the standard *DE/rand/1/bin* algorithm which is best known and mostly used. That is why it is termed as the classic differential evolution algorithm in (Price at al. 2005). Hereafter it is referenced to as CDEA. The new proposed modification of CDEA is termed to as the modified differential algorithm denoted by abbreviation MDEA.

CLASSIC DIFFERENTIAL EVOLUTION ALGORITHM AND GLOBAL CONVERGENCE

In this part we briefly describe the operation of CDEA. Generally, CDEA seeks for the minimum of the cost function by constructing whole generations of

individuals. Each individual is an ordered set of specific values corresponding to one point in the cost function domain. In this way each individual represents a potential solution to the optimization task. The quality of this individual is determined by the evaluation of the cost function corresponding to this individual. The next generation is formed from the existing generation by means of mutation and crossover operators. Specifically, we go successively through all individuals in the generation G. To each individual y_m^G (termed as the *target individual*) we select randomly three other (different) individuals y_{r1}^G, y_{r2}^G, y_{r3}^G from the current generation. We form in a specific way (including randomness) a combination of these three random individuals and the target individual. This combination is termed as the *trial individual* and denoted y_m^{trial}. Then we evaluate the cost function for the target y_m^G and trial individual y_m^{trial} and compare the results. The individual with lower value of the cost function advances to the position of the target individual of the next generation y_m^{G+1}. When this procedure is completed for all target individuals in generation G, we have constructed the new generation of individuals numbered $G + 1$.

The next part illustrates CDEA operation in a definite way in the form of pseudo code.

Input:

Optimization task parameters:

f denotes the cost function, D is the dimension of the cost function domain, $< x_{j\min}, x_{j\max} >$ is a domain of each cost function variable x_j.

CDEA parameters:

NP denotes the generation size (the number of individuals in each generation), NG is the number of calculated generations, F stands for mutation factor ($F \in < 0, 2 >$), and CR denotes the crossover probability ($CR \in < 0, 1 >$). The symbol G stands for the generation number, index m is the number of the individual in the generation, index j describes the j-th component of a specific individual y_m.

Computation:

1. Create an initial generation ($G = 0$) of NP individuals y_m^G, $1 \leq m \leq NP$, (e.g. by use of relation (1)).

2. a) Evaluate all individuals y_m^G of the G-th generation (calculate $f(y_m^G)$ for each individual y_m^G). b) Store the individuals y_m^G and their evaluations $F(y_m^G)$ into matrix **B** (each matrix row contains parameters of individual y_m^G and its

evaluation $F(y_m^G)$. That is matrix be has NP rows and D+1 columns ($1 \leq m \leq NP$).

3. *while* $G \leq NG$

a) *for* $m := 1$ *step* 1 to NP *do*

   (i) randomly select index $s_m \in \{1, 2, \dots, D\}$,

   (ii) randomly select indexes $r_1, r_2, r_3 \in \{1, \dots, NP\}$,

     where $r_l \neq m$ for $1 \leq l \leq 3$;

     $r_1 \neq r_2, r_1 \neq r_3, r_2 \neq r_3$;

   (iii) *for* $j := 1$ *step* 1 to D *do*

     if rand $(0,1) \leq CR$ *or* $j = s_m$) *then*

$$y_{m,j}^{trial} := y_{r_3,j}^G + F\left(y_{r_1,j}^G - y_{r_2,j}^G\right)$$

            else

$$y_{m,j}^{trial} := y_{m,j}^k$$

     end if

     end for (j)

   (iv) *if* $f\left(y_m^{trial}\right) \leq f\left(y_m^k\right)$ *then* $y_m^{G+1} := y_m^{trial}$

          else $y_m^{G+1} := y_m^G$

   end if

  end for (m)

b) store individuals y_m^{G+1} and their evaluations $f\left(y_m^{G+1}\right)$ $(1 \leq m \leq NP)$ of the new $(G+1)$-st generation in the matrix **B**, $G := G + 1$

end while (G).

Output:

The row of matrix **B** that contains the corresponding value $\min\{F(y_m^G); y_m^G \in \mathbf{B}\}$ represents the best found individual y_{opt}.

Comments

The individual y_{opt} in pseudo-code of CDEA is the final solution of the optimization problem.

One way of possible forming the initial generation ($G = 0$) of individuals y_m^0 is given by relation

$$y_{m,j}^0 := x_{j\min} + \text{rand}\,(0, 1) \cdot \left(x_{j\max} - x_{j\min}\right). \quad (1)$$

Values $x_{j\min}$ and $x_{j\max}$ are lower and upper limit of variable x_j. The function rand $(0, 1)$ randomly generates a value from a closed interval $< 0, 1 >$.

Counterexample to Global Convergence of CDEA (Premature Convergence)

It is not difficult to find counterexamples to the global convergence of the CDEA. Let us consider for instance the following two graphs of cost functions with the domain in Euclidean space R^2, see Figure 1. Even for the cost function shown in Figure 1 above the probability that the CDEA finds the global minimum of the cost function is less than one. The reason is that the CDEA

can converge in some cases relatively fast to the local minimum missing completely the global minimum. This results in concentrating the individuals in subsequent generations around the local minimum. As soon as the size of the generation falls under some critical value, the generation is too small to produce trial individuals that could hit the region in the vicinity of the global minimum. This situation is called a *premature convergence*. In this case even increasing the number of generations does not lead to increasing the chance to identify the global minimum. Moreover, the probability that the CDEA finds the global minimum falls with the decreasing measure of the global minimum region. The probability of finding the global minimum for the cost function in Figure 1 below is substantially smaller than for the cost function in Figure 1 above. Additionally, by a sufficient reduction of the measure of the global minimum region this probability can be made as close to zero as possible.
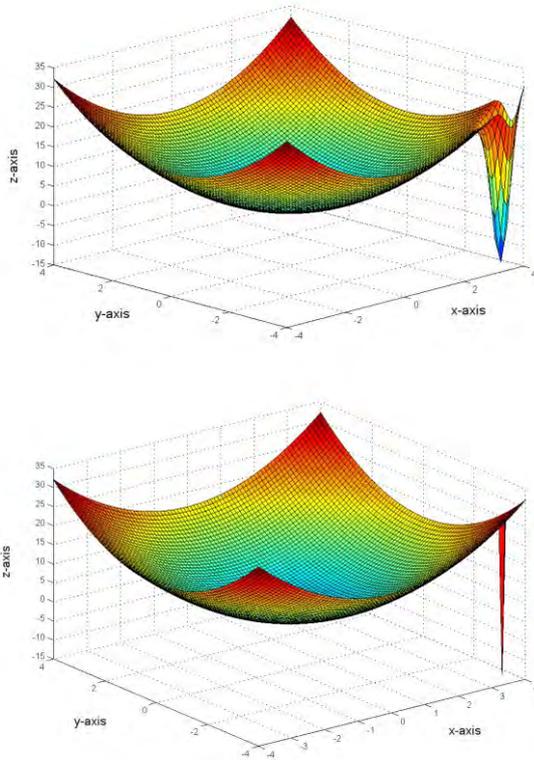
Figure 1: Examples of cost functions with domains in R^2

Numerical Example

We can present a specific cost function to demonstrate the limited ability of CDEA to converge to the global minimum of the cost function. To keep things simple we consider the domain of the cost function as a subset of the two dimensional Euclidean space R^2. We will construct the cost function $F(x_1, x_2)$ as a composition of two simple functions

$$F(x_1, x_2) = F_B(x_1, x_2) + F_M(x_1, x_2). \qquad (2)$$

The term $F_B(x_1, x_2)$ represents the base function. This function is smooth and has one shallow minimum. It can be defined for instance in the following way

$$F_B(x_1, x_2) = x_1^2 + x_2^2,$$

with the domain $D(F_B) = \langle -H, H \rangle \times \langle -H, H \rangle$, where H determines the boundary values of the domain.

The term $F_M(x_1, x_2)$ denotes a modifier function. This function should be relatively steep and with a rather small domain. We use the function $F_M(x_1, x_2)$ to modify the underlying base function $F_B(x_1, x_2)$. The role of the function $F_M(x_1, x_2)$ is to realize the global minimum of the cost function $F(x_1, x_2)$ in relation (2). To be able to construct the function $F_M(x_1, x_2)$ effectively, we introduce another auxiliary function F_P,

$$F_P(x_1, x_2) = x_1^2 + x_2^2 - 1,$$

with the domain $D(F_P) = \{x_1, x_2: x_1^2 + x_2^2 \leq 1\}$. It is obvious that the function $F_P(x_1, x_2)$ is defined exclusively on a unit circle and has values from the closed interval $\langle -1, 0 \rangle$. The graph of the function $F_P(x_1, x_2)$ is a circular paraboloid presented in Figure 2.
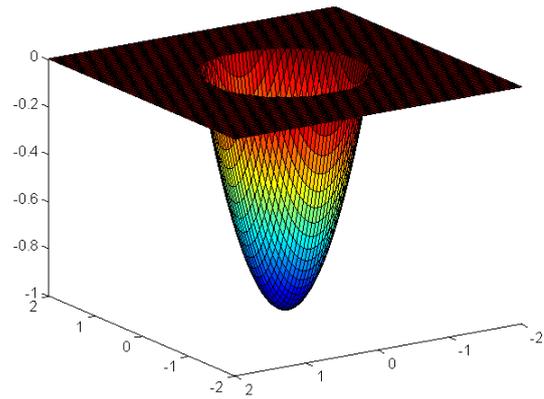

Figure 2: Graph of the auxiliary function $F_P(x_1, x_2)$

The function $F_M(x_1, x_2)$ is then formed as

$$F_M(x_1, x_2) = \lambda_h \cdot F_P\left(\frac{1}{\rho}(x_1 - x_{G1}), \frac{1}{\rho}(x_2 - x_{G2})\right).$$

Here the number λ_h defines the height of the resulting circular paraboloid, ρ denotes the radius of the domain on which the modifier function $F_M(x_1, x_2)$ is defined. Obviously, the modifier function $F_M(x_1, x_2)$ is defined only for points that are closer to the point $[x_{G1}, x_{G2}]$ than the radius of its domain ρ. The coordinates x_{G1}, x_{G2} specify the point, where the modifier function $F_M(x_1, x_2)$ attains its minimum. The overall cost function $F(x_1, x_2)$ is then defined according to the relation (2) by the composite formula

$$F(x_1, x_2) = x_1^2 + x_2^2 + \lambda_h \cdot F_P\left(\frac{1}{\rho}(x_1 - x_{G1}), \frac{1}{\rho}(x_2 - x_{G2})\right)$$

$$(3)$$

We have to choose the parameters λ_h, ρ, x_{G1} and x_{G2} in a reasonable way to obtain the required result. It is clear that we can control the dimensions of the modifier function domain by parameter ρ. The point $[x_{G1}, x_{G2}]$ is to be placed relatively close to the boundary of the cost function domain. This means it is relatively far from the point $[0, 0]$ representing the local minimum of the cost function analogously to the cost functions presented in Figure 1. Since the base function $F_B(x_1, x_2)$ is positive definite, it attains a positive value $F_B(x_{G1}, x_{G2})$ at the point $[x_{G1}, x_{G2}]$. This means we have to take the parameter λ_h sufficiently large, so that the global minimum is essentially lower than the local minimum at the point $[0, 0]$.

We performed numerical experiments with following parameters: $D(F_B) = \langle -H, H \rangle \times \langle -H, H \rangle = \langle -4, 4 \rangle \times \langle -4, 4 \rangle$ with measure $\mu(D\,(F)) = 8^2 = 64$, $x_{G1} = x_{G2} = 3$. This means that the global minimum of the cost function F is at point $[3,3]$ and local minimum at point $[0,0]$. Number of individuals in each generation $NP = 200$, number of generation $NG = 160$, value of parameter $F=0.8$ and value of parameter $CR=0.9$. We realized 200 numerical experiments with value of parameter $\rho = \frac{1}{10}$ (then $\mu(D(F_M)) = \pi\rho^2 \doteq 0.0314$) and 200 numerical experiments with value of parameter $\rho = \frac{1}{16}$ (then $\mu(D(F_M)) \doteq 0.01223$). The results illustrating the limited ability of CDEA to identify the global minimum are summarized in Table 1.

Table 1: Experimental testing of CDEA

CDEA	Local minimum hits	Global minimum hits	Success rate in %
$\rho = \frac{1}{10}$	163	37	18.5
$\rho = \frac{1}{16}$	185	15	7.5

Based on the values given in Table 1, it can be assumed that the decreasing value of ρ (and thus value of $\mu(D(F_M))$) will significantly decrease the success rate of the CDEA algorithm in finding the global minimum.

MODIFIED DIFFERENTIAL EVOLUTION ALGORITHM

As illustrated in the previous part, CDEA does not in general guarantee the convergence to the global minimum of the cost function. This is caused by the too fast convergence of CDEA to the local minimum (premature convergence) resulting in rapid reduction of the generation size (which means a loss of diversity). The most straightforward way how to limit the premature convergence is to replace some individuals with the highest values of the cost function by random individuals in each generation. Though these random individuals reduce partially the convergence speed, they increase substantially the diversity of the generation. In technical terms, it is necessary to make one simple change in the CDEA scheme. We present only the differences with respect to CDEA. See the pseudocode description of CDEA in chapter "Classic Differential Evolution Algorithm and Global Convergence".

Input:
We add another parameter R that determines the ratio of random individuals in each generation, $R \in\,< 0, 1 >$, e.g., $R = 0.1$ means that 10% of individuals in each generation are generated randomly.

Computation:
We add another procedure to the part 3., specifically:
c) determine in matrix B the quantity $\lfloor NP. R \rfloor$ of individuals with the highest cost function values and replace these individuals by randomly generated individuals (e.g. by use of relation (1)) from the search space). Note that here the symbol $\lfloor x \rfloor$ denotes the integer part of the real number x.

This modified algorithm will be called the Modified Differential Evolution Algorithm (MDEA). We applied the numerical experiments on MDEA with the same input parameters as in the previous chapter on CDEA. In addition, the value of R parameter is equal to $R = 0.1$. The results are summarized in Table 2.

Table 2: Experimental testing of MDEA

MDEA	Local minimum hits	Global minimum hits	Success rate in %
$\rho = \frac{1}{10}$	34	166	83.0
$\rho = \frac{1}{16}$	70	130	65.0

Another positive feature of the algorithm MDEA is that if we increase the number of generations NG the global minimum will be identified with an increased probability. This probability can come close to 1 for a sufficiently high number G of generations. We call this aspect of the MDEA an *asymptotic global convergence*. We describe this topic in the following chapter.

ASYMPTOTIC GLOBAL CONVERGENCE OF MDEA

In this part we present several theoretical concepts and statements that can be used to prove the asymptotic global convergence of MDEA. More specifically, we will show that when the number of generations $G \rightarrow \infty$ then the probability that MDEA identifies the global minimum of the cost function approaches 1.

Optimal Solution Set
We would like to find the minimum of the cost function with the lowest value

$$\min\{F(x): x \in S\}, \qquad (4)$$

where S is a measurable search space of a finite measure representing all possible configurations of variables $x = (x_1, x_2, \dots x_n)$. We suppose that the global minimum of function F exists on S. We define a solution set S^* as

$$S^* = \{x^*: F(x^*) = \min\{F(x): x \in S\}\},$$

where x^* represent global minima of the function F. We introduce an optimal solution set S_ε^* as

$$S_\varepsilon^* = \{x \in S: |F(x) - F(x^*)| < \varepsilon\},$$

where $\varepsilon > 0$ is a small positive real number. Denoting by μ the Lebessgue measure, we suppose that for each ε it holds $\mu(S_\varepsilon^*) > 0$.

Convergence in Probability
To examine the global convergence of MDEA we need to introduce a concept of the convergence in probability defined in (Hu et al. 2013).

Definition: Let $\{G(k), k = 1, 2, \dots\}$ be a generation sequence created by a differential evolution algorithm to solve optimization task (4). We say that the algorithm converges to the optimal solution set in probability if

$$\lim_{k \to \infty} p\{G(k) \cap S_\varepsilon^* \neq \emptyset\} = 1, \qquad (5)$$

where p denotes the probability of an event.

Now, we can use this concept to formulate the following statement.

Proposition: Let us suppose that for each generation $G(k)$ of a differential evolution algorithm there exists at least one individual y such that

$$p\{y \in S_\varepsilon^*\} \geq \alpha > 0,$$

where α is a small positive value. Then the algorithm converges to the optimal set S_ε^* in probability. That is relation (5) holds.

The proof of this proposition is stated in full in (Knobloch et al. 2017).

It holds that for each generation G of the MDEA it is true

$$p\{y \in S_\varepsilon^*\} = \alpha \geq 0, \qquad (6)$$

where α is a small positive value. Here $p\{y \in S_\varepsilon^*\}$ denotes the probability that y belongs to S_ε^*. The validity of relation (6) necessarily results from the generation of random individuals in each generation G of MDEA. It follows that MDEA converges to $y \in S_\varepsilon^*$ for any small real positive number ε. This implies the asymptotic global convergence of MDEA. Thus, we know that MDEA converges to the global minimum. The asymptotic convergence of MDEA is proved in detail in (Knobloch et al. 2017), see also (Hu et al. 2013). Probability estimates of reaching the global minimum after performing G generations of MDEA are given in (Knobloch and Mlýnek 2020). These estimates help to decide after how many generations to finish the MDEA calculation.

CONCLUSIONS

CDEA is a universal optimization algorithm that is frequently used in technical projects, economy studies, natural sciences and other important areas of interest. Nevertheless, it has some principal limitations. The main weakness of CDEA is a possible premature convergence of the computing process to a local minimum of the cost function. We demonstrated this fact by means of a simple example.

Identification of this weakness was the starting point for a search of an improved version of the algorithm that would provide better chances regarding the convergence to the global minimum of the cost function. MDEA is a result of these efforts.

MDEA is not prone to the premature convergence because a certain ratio of random individuals in each generation makes it immune to the loss of generation diversity. From the theoretical point of view, we proved that MDEA converges asymptotically to the global minimum of the cost function in probabilistic sense.

The use of MDEA has proved successful to the authors in solving complicated practical optimization problems. For example, it is the task of optimizing the placement of infrared heaters over a metal thin walled mould in the production of artificial leather for the automotive industry (Slush Moulding technology).

The cost function of this optimization problem is a function of many variables (often 300 and more) and has many local minima. Gradient methods, genetic algorithms and also CDEA found only a local minimum of the corresponding cost function (this optimization problem is described in more detail in (Mlýnek and Knobloch 2018) and (Mlýnek et al. 2016). MDEA has also proved successful in optimizing the fibre winding procedures using a fibre-processing head and a non-bearing frame moved by an industrial robot (for more details see (Mlýnek et al. 2020)).

ACKNOWLEDGMENTS

REFERENCES

Affenzeller, M.; Winkler, S.; Wagner, S.; and A. Beham. 2009. "Genetic Algorithms and Genetic Programming." CRC Press, Boca Raton.

Červenka, M. and H. Boudná. 2018. "Visual Guide of F and CR Parameters Influence on Differential Evolution Solution Quality." *Proceedings of 24th International Conference Engineering Mechanics 2018*, Svratka, Czech Republic, 141-144, DOI: 10.21495/91-8-141.

Hu, Z.; S. Xiong; Q. Su; and X. Deng. 2013. "Sufficient Conditions for Global Convergence of Differential Evolution Algorithm." *Journal of Applied Mathematics*, Article ID 139196.

Knobloch, R.; J. Mlýnek; and R. Srb. 2017. "The Classic Differential Evolution Algorithm and Its Convergence Properties". *Applications of Mathematics*, Vol. 62, No. 2, 197-208.

Knobloch, R. and J. Mlýnek. 2020. "Probabilistic Analysis of the Convergence of the Differential Evolution Algorithm." *J. Neural Network World*, Vol. 30, 249-263, DOI: 10.14311/NNW.2020.30.017.

Mlýnek, J.; R. Knobloch; and R. Srb. 2016. "Optimization of a Heat Radiation Intensity and Temperature Field on the Mould Surface." *Proceedings of 30th European Conference on Modelling and Simulation*, Regensburg, Germany, ISBN: 978-0-9932440-2-5, DOI: 10.7148/2016-0425.

Mlýnek, J. and R. Knobloch. 2018. "Model of shell Metal Mould Heating in the Automotive Industry." *Applications of Mathematics*, Vol. 63, No. 2, 111-124.

Mlýnek, J.; M. Petrů; T. Martinec, T. ; and S.S.R. Koloor. 2020. "Fabrication of High-Quality Polymer Composite Frame by a New Method of Fiber Winding Process." *J. Polymers*, Volume 12(5), 30 pages, https://doi.org/10.3390/polym12051037, Open Access.

Price, K.V. 1996. "Differential Evolution - A Fast and Simple Numerical Optimizer." *Proceedings of North American Fuzzy Information Processing. Berkeley*, 524-527.

Price, K.V.; R.M. Storn ; and J.A. Lampien. 2005. "Differential Evolution, A Practical Approach to Global Optimization." Springer-Verlag, Berlin Heidelberg.

Simon, D. 2013. "Evolutionary Optimization Algorithms." John Wiley & Sons, Hoboken, New Jersey.

Storn, R.M. and K.V. Price. 1997. "Differential Evolution - A Simple and Efficient Heuristics for Global Optimization over Continuous Spaces." *Journal of Global Optimization*. Kluwer Academic Publishers, 11, 341-359.

AUTHOR BIOGRAPHIES

ROMAN KNOBLOCH was born in Turnov, the Czech Republic. He finished his studies at the Charles University in Prague, the Faculty of Mathematics and Physics, where he studied physics and teaching of mathematics and physics. His main areas of interest are: modelling of physical phenomena, modern optimization methods and heat and transport phenomena in continuum mechanics. He works as an assistant professor at the Technical University of Liberec where he also graduated his PhD study programme. His e-mail address is roman.knobloch@tul.cz

JAROSLAV MLÝNEK was born in Trnava, Czechoslovakia and went to the Charles University in Prague, where he studied numerical mathematics at the Faculty of Mathematics and Physics and he graduated in 1981. In his work he focuses on the computational problems of heating and thermal losses in components of electrical machines and on optimization procedures. Currently he works as an associate professor at the Technical University of Liberec. His e-mail address is: jaroslav.mlynek@tul.cz

COMPARATIVE EVALUATION OF *LACTOBACILLUS PLANTARUM* STRAINS THROUGH MICROBIAL GROWTH KINETICS

Georgi Kostov*, Rositsa Denkova-Kostova**, Vesela Shopska*, Bogdan Goranov***, Zapryana Denkova***

*Department of Wine and Beer ** Department of Biochemistry and Molecular Biology, *** Department of Microbiology

University of Food Technologies, 4002, 26 Maritza Blvd., Plovdiv, Bulgaria

E-mail: george_kostov2@abv.bg; rositsa_denkova@mail.bg; vesi_nevelinova@abv.bg; goranov_chemistry@abv.bg; zdenkova@abv.bg

KEYWORDS

Probiotics, growth kinetics, modeling, optimization

ABSTRACT

The study of the growth kinetics of lactobacilli with pronounced probiotic properties in their batch cultivation is essential. Various models based on the logistic curve model, containing parameters showing the influence of the accumulating lactic acid on the biosynthesis of the product, as well as parameters showing the sensitivity of the cells to lactic acid were used to model the growth kinetics in the present work. The rate constant of adaptation of the studied strains to the used nutrient medium and the induction period were also determined. The kinetics of lactic acid synthesis was determined according to the Weibull model.

INTRODUCTION

The bacteria most commonly included as components of probiotic preparations are lactic acid bacteria (*Lactobacillus* sp.*, Enterococcus* sp.*, Pediococcus* sp.*, Streptococcus* sp.*, Lactococcus* sp.*, Leuconostoc* sp.) and bifidobacteria. They are also used in the production of probiotic foods (Gibson, 2004), with the largest share being that of the lactobacilli.

Not all species of lactobacilli, as well as not all strains of the same species can be included in the composition of probiotics, but only those that have certain properties (Saarela et al., 2002): to be part of the natural microflora in humans and animals; to be able to adhere and colonize the intestinal mucosa to compete with enteropathogenic bacteria for adhesion sites and nutrients; to survive and maintain their activity in the conditions of the gastrointestinal tract; to be able to reproduce in the gastrointestinal tract; to have high antimicrobial activity in order to suppress and expel pathogenic and toxigenic microorganisms from the biological niche; to allow industrial cultivation - to maintain their activity during production and storage; to modulate the immune response and to be safe for clinical and nutritional use. *Lactobacillus plantarum* is a flexible and versatile species of lactic acid bacteria, which is often found in many probiotic, functional and fermented foods and beverages (cheeses, fermented milk, pasta, sausages and various vegetable juices) or is used as a probiotic (Gobbetti et . al., 1994a; Gobbetti et al., 1994b; Corsetti and Gobbetti, 2002; Guidone et al., 2014). This is due to its flexible metabolism, its ability to adapt to different environmental conditions and the wide range of antimicrobial activity it possesses (Di Cagno et al., 2009).

Along with its antimicrobial activity, the active cells of *L. plantarum* 13M5 have the ability to destroy the mycotoxin patulin at a concentration of 5 mg/dm^3 as a result of the synthesis of a bacteriocin called plantaricin (Todorov et al., 1999; Wei et al., 2020). *Lactobacillus plantarum* YJ7 shows antihyperglycemic potential and reduces insulin resistance, so it can be used in the composition of drugs targeted at people suffering from diabetes (Zhong et al., 2020). In experimental animals, *Lactobacillus plantarum* strains, and in particular *Lactobacillus plantarum* LP33, have been shown to reduce liver damage due to lead intoxication (Hu et al., 2020).

The main metabolite of lactic acid fermentation is lactic acid. It is known that its increasing concentration during fermentation has an inhibitory effect on the growth of the microbial population. The sensitivity to the accumulating lactic acid is strain-specific (Bouguettoucha et al., 2011; Gordeev et al., 2017). The selected mathematical models contain parameters characterizing the influence of lactic acid on lactobacilli. It is also important to determine both the induction period - the time from the lag phase, during which the cells begin to synthesize the necessary cellular structures and enzymes and to move from unadapted to adapted state to the composition of the medium and culture conditions, and the rate constant of adaptation (Warfolomeev and Gurevich, 1999). One of the important conditions for comparing the kinetic characteristics of the models is the initial conditions - inoculum and acidity of the medium – to be the same. Since this is difficult to achieve, it is necessary to measure the data of the biomass and the titratable acidity in the models (Tishin and Fedorov, 2016; Tishin and Golovinskaya, 2015). As a result of the above-mentioned features, the following mathematical models were chosen to model the kinetics of growth and acid formation:

$$\frac{dX_b}{d\tau} = \mu_{max} \left(1 - \frac{P_b}{P_{bm}}\right)^c X_b \qquad (1)$$

$$\frac{dX_b}{d\tau} = \mu_{max} \left(1 - \frac{X_b}{X_{bm}}\right)^n X_b \qquad (2)$$

$$\frac{dX_b}{d\tau} = \mu_{max} \left[\left(1 - \frac{X_b}{X_{bm}}\right)^{n_1} \left(1 - \frac{P_b}{P_{bm}}\right)^q\right] X_b \qquad (3)$$

$$\ln \frac{M}{N_0} = \mu\tau + \ln\left\{\frac{k_0}{k_0 + \mu}\left[1 + \frac{\mu}{k_0} e^{[-(k_0+\mu)\tau]}\right]\right\} \qquad (4)$$

$$K_T = a - be^{-(q_p\tau)^\delta} \qquad (5)$$

where: μ_{max} - maximum specific growth rate, h^{-1}; X_b, P_b, X_{bm} and P_{bm} are the biomass, the lactic acid amount, the final concentration of the biomass and the lactic acid, respectively, in dimensionless form; M - current

biomass concentration, cfu/cm³; N_0 - initial biomass concentration, cfu/cm³; τ_a - induction period, h; k_0 - rate constant of cell adaptation to the medium and culturing conditions, h⁻¹; c - a parameter taking into account the inhibitory effect of the accumulating product (lactic acid) on the cell growth; n and n_1 - coefficients taking into account the influence of lactic acid on the cells, respectively showing the resistance (sensitivity) of the cells to the increasing concentration of the product; q - coefficient showing the inhibitory effect of the product, lactic acid, on its own synthesis; K_T - titratable acidity in dimensionless form; a - maximum value of the titratable acidity in dimensionless form; b - coefficient equal to the difference between the maximum and initial titratable acidity in dimensionless form; q_P - specific rate of acid formation, h⁻¹; δ - an indicator determining the change in the shape of the curve or the change in the rate of accumulation of lactic acid over time; τ - cultivation time, h.

The presented models make it possible to determine the parameters of the fermentation process analytically. Moreover, they allow the assession of the influence of cultivation conditions and the accumulation of lactic acid on the microbial population.

The aim of the present work was to study the lactic acid fermentation process with selected probiotic *Lactobacillus plantarum* strains by applying modified dependences of the logistic curve type and assessing the influence of acid formation on the lactic acid fermentation process.

MATERIALS AND METHODS

Microorganisms and cultivation conditions

The study was conducted with four different strains of *Lactobacillus plantarum*: *Lactobacillus plantarum* 4/17, *Lactobacillus plantarum* 3, *Lactobacillus plantarum* 10 and *Lactobacillus plantarum* 1/18, isolated from spontaneously fermented vegetables. The 4 strains were identified by molecular-genetic identification method – 16S rDNA sequencing – as representatives of the *Lactobacillus plantarum* species (unpublished data).

Cell cultivation was performed under static conditions in flasks using LAPTg10-broth medium. Samples were periodically taken to determine the titratable acidity of the medium and the number of viable lactobacilli cells.

Nutrient media

• LAPTg10-broth;
• MRS-agar;
• Saline solution.

Methods of analysis

• Determination of titratable acidity (ISO/TS 11869:2012);
• Number of viable lactobacilli cells (ISO 7889:2005).

Identification of the model parameters

The logistic curve models from 1 to 3 are solved numerically using the Runge-Kuta method of the 4ᵗʰ row, and the parametric identification in them is performed by minimizing the sum of the squares of the difference between the experimental data and the data obtained from the corresponding model in Microsoft Excel (Choi et al., 2014). The parametric identification of model 4 and the Weibull model was performed using the software Curve Expert Professional by nonlinear regression.

RESULTS AND DISCUSSION

Table 1 presents the data from the determination of the induction period and the rate constant of adaptation, determined according to equation 4.

Table 1: Induction period and rate constant of adaptation

Strain	τ_a, h	k_0, h⁻¹
L. plantarum 4/17	0.36	0.256
L. plantarum 3	0.73	0.253
L. plantarum 10	0.88	0.227
L. plantarum 1/18	1.43	0.103

The strains *L. plantarum* 4/17 and *L. plantarum* 3 have a significantly shorter induction period (0.36 h and 0.73 h, respectively) and higher values of the rate constant of adaptation (0.256 h⁻¹ and 0.253 h⁻¹, respectively), compared to the other two strains studied (Table 1). The longest induction period of 1.43 h and the lowest rate constant of adaptation (0.103 h⁻¹) was observed for *L. plantarum* 1/18, while *L. plantarum* 10 occupied an intermediate place with an induction period of 0.88 h and a rate constant of adaptation of 0.227 h⁻¹.

From the studies conducted it can be concluded that *L. plantarum* 4/17 would most quickly adapt to the fermentation medium and cultivation conditions, followed by *L. plantarum* 3 and *L. plantarum* 1/18. This shows that the fermentation medium used for these strains has an optimal composition and is suitable for their growth. The slower adaptation of *L. plantarum* 10 compared to other strains may be due to the lack of some substrates in the medium, values of the redox potential and temperature regime different from the optimal ones for the specific strain studied. Another reason for the lower values of the rate constant of adaptation and the longer induction period is probably the static nature of the medium, which is characterized by a lack of surface aeration, which for some species of *L. plantarum* has certain stimulating effect, as many strains of this species are microaerophiles.

Table 2A presents the kinetic parameters of the used logistic curve models (equation 1 to equation 3). Table 2B presents the correlation coefficients and errors of the models. The comparison of the models with the experimental data is presented in Fig. 1 to Fig. 4. As a general conclusion, the three models used describe the experimental data from the cultivation of the four *L. plantarum* strains with very high accuracy. Similarly, the model used to describe the kinetics of lactic acid accumulation describes very well the experimental data (equation 5). This general conclusion can be explained by the fact that, unlike numerical methods for determining the kinetic parameters in the analytical solution of differential equations, the number of parameters in the model is minimized, and the obtained parameters have a

clear biological meaning. It is this biological meaning that we will demonstrate by interpreting the results for the four strains studied. Another reason for increasing the accuracy of the models is the dimensionless form of the biomass, which reduces the identification error. In this way, the influence of dimensionality and the influence of random errors in the enumeration of microorganisms according to the methodology for determining the concentration of viable cells is avoided.

Table 2A: Kinetic parameters in the different logistic curve models in the cultivation of the *Lactobacillus plantarum* strains

Strain	Mathematical model										
	Model 1 (eq.1)			Model 2 (eq.2)			Model 3 (eq.3)				
	μ_m, h^{-1}	c	P_m	μ_m, h^{-1}	n	X_m	μ_m, h^{-1}	n_1	q	P_m	X_m
4/17	0.0198	0.346	1.98	0.0210	3.661	2.41	0.0181	3.208	0.117	2.81	2.30
3	0.0190	0.329	2.48	0.0476	1.703	2.83	0.0383	1.257	0.116	2.78	2.54
10	0.0220	0.368	2.38	0.0740	1.135	2.16	0.0400	0.980	0.130	2.85	2.45
1/18	0.0190	0.448	2.83	0.0156	0.857	1.87	0.0470	0.478	0.621	1.91	2.64

Table 2B: Correlation coefficients and errors

Strain	Mathematical model					
	Model 1 (eq.1)		Model 2 (eq.2)		Model 3 (eq.3)	
	R^2	Error	R^2	Error	R^2	Error
4/17	0.9364	0.078	0.9405	0,078	0.9474	0.074
3	0.9115	0.174	0.9378	0.160	0.9461	0.159
10	0.9022	0.135	0.8200	0.155	0.9406	0.159
1/18	0.8853	0.153	0.9497	0.122	0.9495	0.123
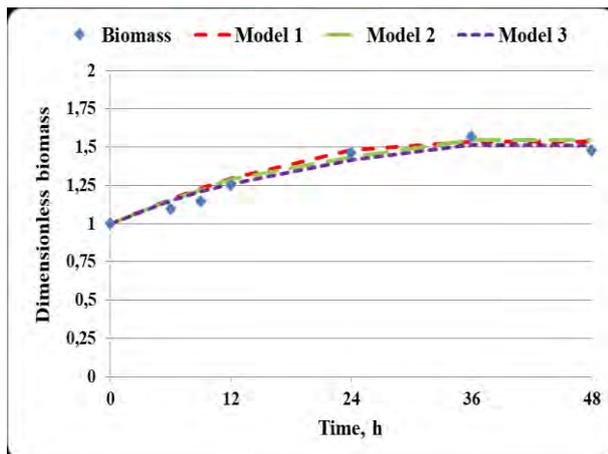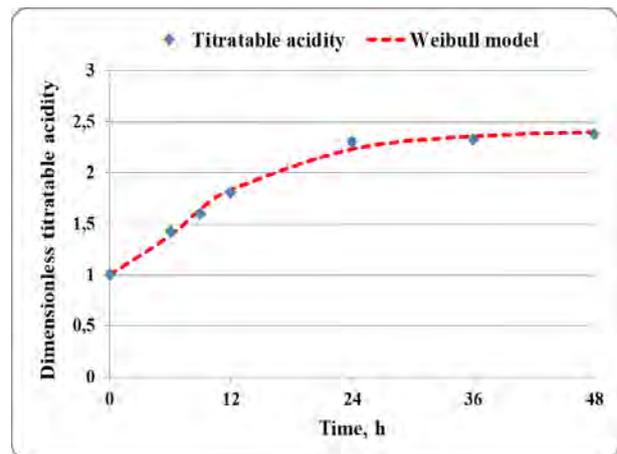
a) biomass   b) lactic acid

Figure 1: Biomass and lactic acid accumulation kinetics for *Lactobacillus plantarum* 4/17
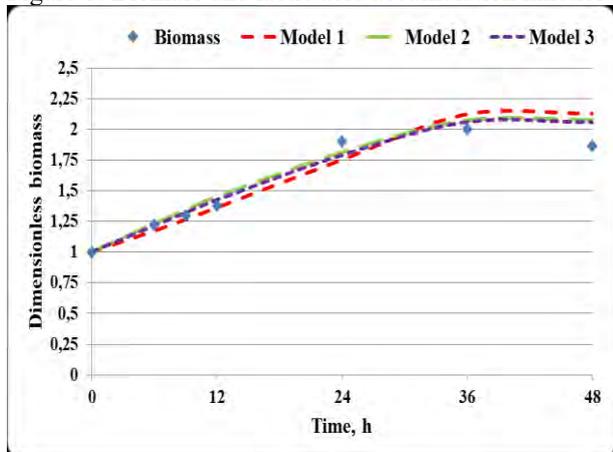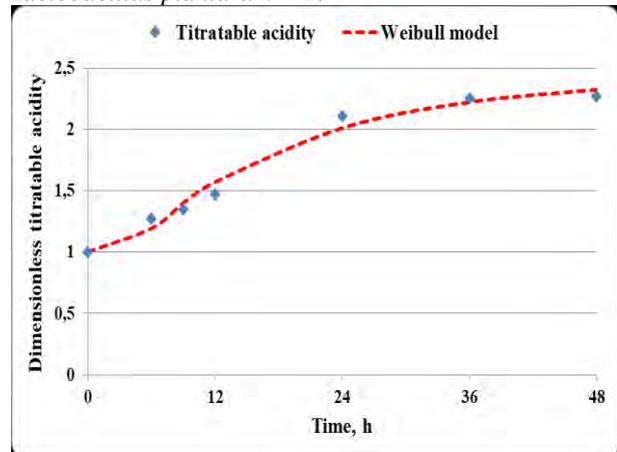
a) biomass   b) lactic acid

Figure 2: Biomass and lactic acid accumulation kinetics for *Lactobacillus plantarum* 3

In all four studied strains the increasing concentration of lactic acid would have a less pronounced inhibitory effect on the maximum specific growth rate (Table 2). This is underlined by the relatively low values of the parameter *c* in model 1. In *L. plantarum* 4/17, *L. plantarum* 3 and *L. plantarum* 10 this parameter has comparable values - 0.346, 0.329 and 0.368, respectively. A higher value of the parameter *c* is observed in *L. plantarum* 1/18 - 0.448. The observed higher value of the parameter can be explained by the higher concentration of lactic acid accumulated by the strain. This is evidenced by the value of the parameter **Pm**, which is 2.83 and its value is the highest one among the **Pm** values in all the four *L. plantarum* strains studied. The lowest final concentration of lactic acid in dimensionless form is observed for *L.*

plantarum 4/17 - 1.98, while for *L. plantarum* 3 and *L. plantarum* 10 it has comparable values - 2.48 and 2.38, respectively. According to the data from model 1, *L. plantarum* 10 has the highest maximum specific growth rate of 0.0220 h^{-1}. The remaining three strains are characterized by lower and commensurable maximum specific growth rates, varying in the range from 0.0190 h^{-1} to 0.0198 h^{-1}.

It is interesting to determine the effect of lactic acid on the cultivation process. Model 2 in which the parameter **n** is subjected to identification is used to achieve this goal. This parameter shows the effect of lactic acid on the biomass, or, more precisely, the resistance of the cells to the accumulating lactic acid. The data are summarized in Table 2.
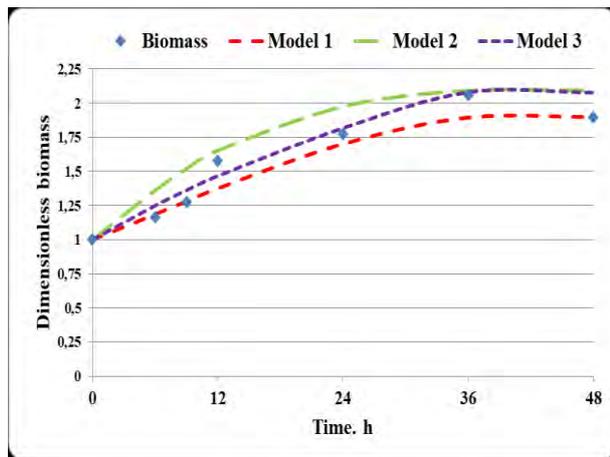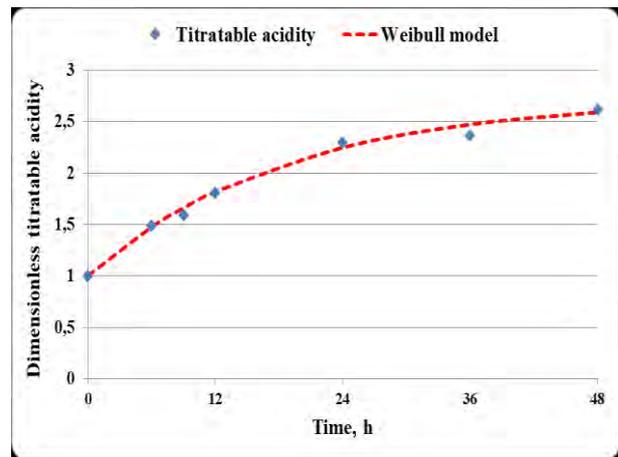
a) biomass   b) lactic acid

Figure 3: Biomass and lactic acid accumulation kinetics for *Lactobacillus plantarum* 10
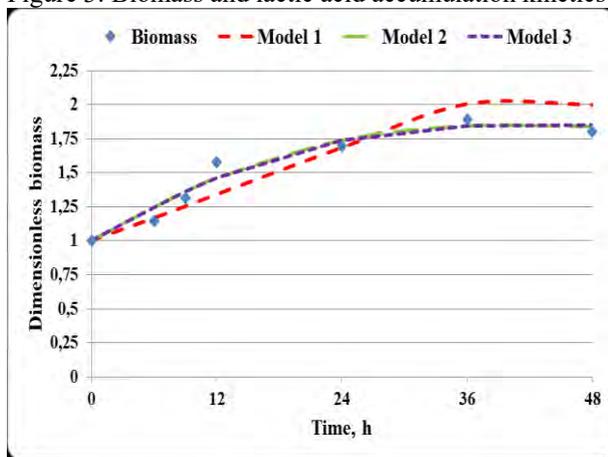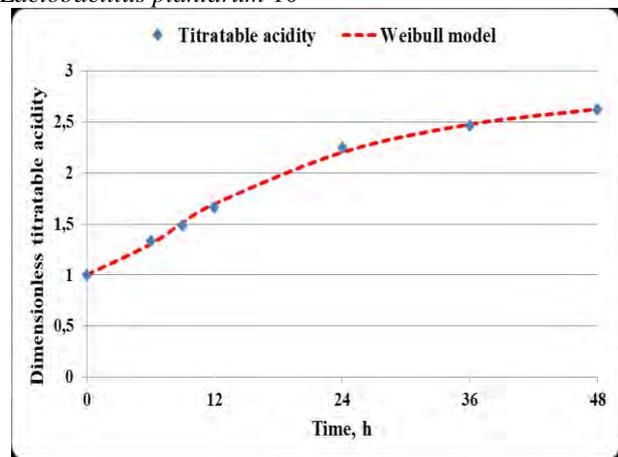
a) biomass   b) lactic acid

Figure 4: Biomass and lactic acid accumulation kinetics for *Lactobacillus plantarum* 1-18

The cells of *L. plantarum* 4/17 (Table 2) show the highest sensitivity, respectively the lowest resistance to the increasing concentration of lactic acid. For this strain the parameter **n** is 3.661. *L. plantarum* 1/18 shows the lowest sensitivity to the increasing concentration of lactic acid (n = 0.857). *L. plantarum* 3 and *L. plantarum* 10 have intermediate resistance to the metabolic product, with the values of **n** being 1,703 and 1,135, respectively. According to model 2, the highest maximum specific growth rate was observed for *L. plantarum* 10 (***μₘ*** =

0.0740 h^{-1}), followed by *L. plantarum* 3 (***μₘ = 0.0470 h^{-1}***). *L. plantarum* 4/17 and *L. plantarum* 1/18 have lower values of the maximum specific growth rate - ***μₘ = 0.0210 h^{-1}*** and ***μₘ = 0.0156 h^{-1}***, respectively.

According to the data from model 2, the highest final concentration of biomass in dimensionless form is achieved by *L. plantarum* 3 - 2.83. In *L. plantarum* 4/17 and *L. plantarum* 10 the maximum concentration of biomass in dimensionless form is 2.41 and 2.16, respectively, and in *L. plantarum* 1/18 it is 1.87.

The meaning of parameter n_1 in Equation 3 is similar. According to this model, *L. plantarum* 4/17 is again characterized by the highest sensitivity to the accumulating lactic acid. In this strain n_1 is 3.208. The next strain in terms of cell sensitivity to the increasing concentration of lactic acid is *L. plantarum* 3, which is also characterized by a high value of the parameter n_1 (1.257). However, in *L. plantarum* 10 model 3 predicted a value of the parameter n_1 below 1, namely 0.980 (the value of the analogous parameter n in model 2 is 1.135). This shows that according to model 3, the strain is characterized by increased resistance (reduced sensitivity) to the increasing concentration of lactic acid. The highest resistance to lactic acid is demonstrated by *L. plantarum* 1/18, characterized by the lowest value of the parameter n_1 (0.478) (Table 2A).

The comparative assessment of the resistance of different *Lactobacillus plantarum* strains to lactic acid is important for the characterization of the strains. In general, it can be assumed that strains that have higher resistance to lactic acid would also have a higher survival rate at low pH in the human gastrointestinal tract. This is especially important for the selection of strains, to be included in the composition of probiotic preparations, because resistance to low pH is one of the most important requirements to potentially probiotic strains.

An important parameter in model 3 is the parameter q. It reflects the inhibitory effect of lactic acid on its synthesis rate. In *L. plantarum* 4/17, *L. plantarum* 3 and *L. plantarum* 10 this parameter has low values - 0.117, 0.116 and 0.130, respectively (Table 2). This in turn shows high intensity of acid formation and a relatively weak inhibitory effect of the acid on its synthesis rate.

In contrast, in *L. plantarum* 1/18 the value of q is 0.621. This means that the increasing acid concentration would have stronger inhibitory effect on the lactic acid synthesis rate. The high value of q characterizes this strain with lower energy of acid formation and therefore the process of acid formation would be more moderate, and the least amount of lactic acid in dimensionless form is accumulated in the medium - 1.91.

The other three strains are characterized by high values of the final concentration of lactic acid in dimensionless form – from 2.78 to 2.85. Unlike model 2, here the values of the biomass in dimensionless form at the end of the fermentation process vary in a relatively small range - from 2.30 to 2.64 (Table 2A).

In order to confirm the assumptions about the acid-forming ability of the studied strains, the specific rate of acid formation q_P and the degree of change in the intensity of lactic acid accumulation over time in general (δ) were calculated. The results of the conducted modeling are presented in Table 3 and Table 4 and the models for the studied strains are shown in real form.

The results in Table 3 once again confirm the conclusions made about the acid-forming ability of the studied strains. According to the Weibull model, *L. plantarum* 4/17, *L. plantarum* 3 and *L. plantarum* 10 have high values of the parameter δ - 2.55, 1.55 and 1.88, respectively. This indicates that in these strains the acid-formation process would proceed with a higher intensity over time in general, compared to *L. plantarum* 1/18. In *L. plantarum* 1/18 δ has a value less than 1, namely 0.99, which again confirms that in this strain the acid formation would be more moderate in time as a whole, although in this strain the Weibull model predicts a slightly higher rate of lactic acid synthesis (0.054 h^{-1}) compared to the other strains. In the other strains, the specific rate of acid formation occupies close values and varies in the range from 0.027 h^{-1} to 0.054 h^{-1} for the different strains.

Table 3: Kinetic parameters in the Weibull model in the cultivation of the *Lactobacillus plantarum* strains

Strain	a	b	q_P, h^{-1}	Δ	R^2	Error
4/17	2.42	1.82	0.037	2.55	0.9933	0.20
3	2.42	1.77	0.041	1.51	0.9803	0.20
10	2.81	2.04	0.027	1.88	0.9981	0.23
1/18	2.72	1.72	0.054	0.99	0.9900	0.22

Table 4: Weibull's mathematical models in real form

Strain	Models in real form
L. plantarum 4/17	$K_T = 2{,}42 - 1{,}82e^{(-0{,}032\tau)^{2{,}55}}$
L. plantarum 3	$K_T = 2{,}42 - 1{,}77e^{(-0{,}047\tau)^{1{,}51}}$
L. plantarum 10	$K_T = 2{,}81 - 2{,}04e^{(-0{,}027\tau)^{1{,}88}}$
L. plantarum 1/18	$K_T = 2{,}72 - 1{,}72e^{(-0{,}054\tau)^{0{,}99}}$

CONCLUSION

Some important conclusions can be drawn for the modeling of the fermentation processes and in particular lactic acid fermentation, from the obtained results. The data show that the use of only one kinetic model does not show all aspects of the lactic acid fermentation process.

Combining several mathematical dependencies makes it possible to consider different aspects of the process. For example, equation 4 allows the estimation of the time for adaptation of the culture and the possibility for the real process to start faster. Equations 1 to 3 make it possible to assess the various aspects of the fermentation process - the accumulation of biomass, the influence of lactic

acid, both on the biomass growth and on the acid-formation rate.

The possibility of the models used to assess the sensitivity of the strains to their own metabolic product allows the selection of high-resistant strains to be used in the composition of probiotic preparations, but also the selection of strains that produce less lactic acid and can be used in food development and production.

Therefore, the modeling of the fermentation process must be done with at least two dependencies that reflect the different aspects of the modeled process. Thus, is it possible to achieve a complete interpretation of the various aspects of fermentation. In addition, the dependencies proposed in the present paper allow the estimation of kinetic parameters to be done through simple analytical dependencies. This allows for faster process management decisions.

The main purpose of the present study was to allow the evaluation of different strains of lactic acid bacteria with a view to their use in the production of different types of functional foods. Knowledge of the fermentation kinetics and the behavior of the strains under different cultivation conditions makes it possible to model the fermentation process, and hence the composition of the obtained functional foods.

In this regard, the results allow the strains to be divided into two groups - strains with high growth rate (strain 10), strains with moderate growth rate and high rate of acid formation (strain 1/18) and strains with moderate growth rate and moderate acid formation (strains 4/17 and 3). Depending on the specific food production, the choice may fall on different groups of strains. In some cases, the functional characteristics of a specific food product are determined by the high concentration of viable cells, while in other cases - by the lactic acid produced by the lactic acid bacteria strains and, hence, accumulated in the food product. In this sense, the combination of different models to describe the kinetics of microbial growth allows for improved options for selection and management of the process of functional food production.

REFERENCES

Bouguettoucha, A., B. Balannec and A. Amrane. 2011. "Unstructured Models for Lactic Acid Fermentation: A Review." *Food Technol. Biotechnol.*, 49 (1), 3–12.

Choi, M., M. Saeed Al-Zahrani, and S. Y. Lee. 2014. "Kinetic model-based feed-forward controlled fed-batch fermentation of *Lactobacillus rhamnosus* for the production of lactic acid from Arabic date juice." *Bioprocess Biosyst Eng*., 37, 1007–1015. https://doi.org/10.1007/s00449-013-1071-7

Corsetti, A. and M. Gobbetti. 2002. "*Lactobacillus plantarum*". In *Encyclopedia of dairy sciences* (H. Proginsli, J.W. Fuquay, and P.F. Fox Eds.), New York: Academic Press Ltd., 1501-1507.

Di Cagno, R., M. De Angelis, R. Coda, F. Minervini, and M. Gobbetti. 2009. "Molecular adaptation of sourdough *Lactobacillus plantarum* DC400 under co-cultivation with other lactobacilli." *Research in Microbiology,* 160, 358-366. https://doi.org/10.1016/j.resmic.2009.04.006

Gibson, G. R. 2004. "From probiotics to prebiotics and a healthy digestive system". *J. Food Science*, 69 (5), M141- M143. https://doi.org/10.1111/j.1365-2621.2004.tb10724.x

Gobbetti, M., A. Corsetti, and J. Rossi. 1994b. "The sourdough microflora, evolution of soluble carbohydrates during the sourdough fermentation." *Microbiologie Aliments Nutrition*, 12, 9–15.

Gobbetti, M., A. Corsetti, J. Rossi, F. La Rosa, and M. De Vincenzi. 1994a. "Identification and clustering of lactic acid bacteria and yeasts from wheat sourdoughs of central Italy." *Ital J Food Sci*, 1, 85–93.

Gordeev, L., A. Koznov, A. Skichko, and Y. Gordeeva. 2017. "Unstructured mathematical models of the lactic acid biosynthesis kinetics: A Review." *Theoretical Foundations of Chemical Engineering*, 51 (2), 175-190.

Guidone, A., T. Zotta, R. P. Ross, C. Stanton, M. Rea, E. Parente, and A. Ricciardi. 2014. "Functional properties of *Lactobacillus plantarum* strains: A multivariate screening study." *LWT - Food Science and Technology*, 56, 69-76. https://doi.org/10.1016/j.lwt.2013.10.036

Hu, T., J. Song, W. Zeng, J. Li, H. Wang, Y. Zhang, and H. Suo. 2020. "*Lactobacillus plantarum* LP33 attenuates Pb-induced hepatic injury in rats by reducing oxidative stress and inflammation and promoting Pb excretion." *Food and Chemical Toxicology*, 143. Paper ID: 111533. https://doi.org/10.1016/j.fct.2020.111533

ISO/TS 11869:2012. Fermented milks — Determination of titratable acidity — Potentiometric method

ISO 7889:2005. Yogurt — Enumeration of characteristic microorganisms — Colony-count technique at 37 degrees C

Saarela, M., L. Zahteenmaki, R. Crittenden, S. Salminen, and T. Mattila-Sandholm. 2002. "Gut bacteria and health foods – the European perspective." *Int. J. Food Microbiol*. 78, 99-117.

Tishin, V. B., and A. V. Fedorov. 2016. The peculiarities of mathematical modelling for the kinetics of microorganisms' cultivation." *Processes and Food Production Equipment*, 9 (4), 65-74. (in Russian).

Tishin, V. B., and O. V. Golovinskaia. 2015. *Experiment search and mathematical models of the kinetics of biological processes. Textbook*. St. Petersburg, University ITMO Publ., p. 111. (in Russian)

Todorov, S., B. Onno, O. Sorokine, J. M. Chobert, I. Ivanova, X. Dousset. 1999. "Detection and characterization of a novel antibacterial substance produced by *Lactobacillus plantarum* ST31 isolated from sourdough." *Int. J. Food Microbiol*., 48, 167–177. https://doi.org/10.1016/S0168-1605(99)00048-3

Warpholomeew, S. D and K. G. Gurevich. 1999. *Biokinetic –practical course*", Fair-Press, Moscow, ISBN: 5-8183-0050-1, p.720. (in Russian).

Wei, C., L. Yu, N. Qiao, S. Wang, F. Tian, J. Zhao, H. Zhang, Q. Zhai, and W. Chen. 2020. "The characteristics of patulin detoxification by *Lactobacillus plantarum* 13M5." *Food and Chemical Toxicology*, 146, Paper ID 111787. https://doi.org/10.1016/j.fct.2020.111787

Yoha, K.S., J. A. Moses, and C. Anandharamakrishnan. 2020. "Effect of encapsulation methods on the physicochemical properties and the stability of *Lactobacillus plantarum* (NCIM 2083) in synbiotic powders and *in-vitro* digestion conditions." *J. Food Eng.*, 283, Paper ID: 110033. https://doi.org/10.1016/j.jfoodeng.2020.110033

Zhong, H. Abdullah, Y. Zhang, M. Zhao, J. Zhang, H. Zhang, Y. Xi, H. Cai, and F. Feng. 2020. "Screening of novel potential antidiabetic *Lactobacillus plantarum* strains based on *in vitro* and *in vivo* investigations." *LWT – Food science and technology*, 139, Paper ID: 110526. https://doi.org/10.1016/j.lwt.2020.110526

ACKNOWLEDGEMENTS

AUTHOR BIOGRAPHIES

GEORGI KOSTOV is Professor at the Department of Wine and Beer Technology at the University of Food Technologies, Plovdiv. He received his MSc degree in Mechanical Engineering in 2007, a PhD degreein Mechanical Engineering in the Food and Flavor Industry (Technological Equipment in the Biotechnology Industry) in 2007 at the University of Food Technologies, Plovdiv, and holds a DSc degree in Intensification of Fermentation Processes with Immobilized Biocatalysts. His research interests are in the area of bioreactor construction, biotechnology, microbial population investigation and modeling, hydrodynamics and mass transfer problems, fermentation kinetics, and beer production.

VESELA SHOPSKA is Head Assistant Professor at the Department of Wine and Beer Technology at the University of Food Technologies, Plovdiv. She received her MSc degree in Wine-making and Brewing Technology in 2006 at the University of Food Technologies, Plovdiv. She received her PhD in Technology of Alcoholic and Non-alcoholic Beverages (Brewing Technology) in 2014. Her research interests are in the area of beer fermentation with free and immobilized cells, yeast and bacteria metabolism and fermentation activity.

ROSITSA DENKOVA-KOSTOVA is Head Assistant Professor at the Department of Biochemistry and Molecular Biology at the University of Food Technologies, Plovdiv. She received her MSc degree in Industrial Biotechnologies in 2011 and a PhD degree in Biotechnology (Technology of Biologically Active Substances) in 2014. Her research interests are in the area of isolation, biochemical and molecular-genetic identification and selection of probiotic strains and development of starters for functional foods.

BOGDAN GORANOV is an assistant at the department of Microbiology at the University of Food Technologies, Plovdiv. He received his PhD in 2015 from the University of Food Technologies, Plovdiv. The theme of his thesis was "Production of Lactic Acid with Free and Immobilized Lactic Acid Bacteria and its Application in the Food Industry". His research interests are in the area of bioreactor construction, biotechnology, microbial population investigation and modeling, hydrodynamics and mass transfer problems, and fermentation kinetics.

ZAPRYANA DENKOVA is a professor at the department of Microbiology at the University of Food Technologies, Plovdiv. She received her MSc in "Technology of microbial products" in 1982, PhD in „Technology of biologically active substances" in 1994 and DSc on "Production and application of probiotics" in 2006. Her research interests are in the area of selection of probiotic strains and development of starters for food production, genetics of microorganisms, and development of functional foods.

This publication has been specially selected for reprinting by the "Simulation and Optimization" track of the 35th ECMS International Conference on Modelling and Simulation.

MAKE-TO-ORDER PRODUCTION PLANNING WITH SEASONAL SUPPLY IN CANNED PINEAPPLE INDUSTRY

Kanapath Plangsrisakul
Tuanjai Somboonwiwat
Chareonchai Khompatraporn
Department of Production Engineering,
King Mongkut's University of Technology Thonburi (KMUTT), Bangkok 10140 Thailand
E-mails: kanapath.002@mail.kmutt.ac.th, tuanjai.som@kmutt.ac.th, charoenchai.kho@kmutt.ac.th

KEYWORDS

Canned Pineapple Industry, Make-to-order Inventory, Multi-products Multi-periods Production Planning, Pineapple Color Ratios, Seasonal Raw Materials.

ABSTRACT

This research studies a make-to-order production planning problem in a canned pineapple industry. Pineapple is a seasonal perishable fruit. Thus, the cost of fresh pineapple which is the main raw material in canned pineapple products is inexpensive during its season because of its abandance. The color of the pineapple also determines the price of the canned pineapple. However, the availability of different colors (referred as "choice" and "standard") is dependent. Specifically, if for a given month the ratio of the choice-color pineapple increases, the ratio of the standard-color pineapple decreases. There are several costs involve such as fresh pineapple cost, can cost, sugar cost, water cost, labor cost, energy cost, and inventory holding cost. This problem is formulated as a mathematical model to maximize the total profit over four-months planning horizon. Two supply uncertainty cases are tested which are low and high ratios of the choice color. The results show that the profit depends on available color ratios of the pineapple. The production planning is best if it matches with the availability of the color ratios. In certain months, some fresh pineapple purchased exceed the need of the production because of the dependency of the two colors. The inventory holding cost also influences the production decision—whether to produce the canned pineapple in earlier months or it is better to produce only the canned pineapple when it is needed to serve the customer orders.

INTRODUCTION

Thailand is the global exporter of canned pineapple with the market share of 37.2% worldwide or USD 338.09 million in value, followed by the Philippines and Indonesia (based on the 2019 statistics) (TRIDGE, 2019). The three countries together cover about 70% of the world's market share (Wattanakul et al., 2020). Pineapple is a seasonal fruit but the demands for canned pineapple exist throughout the year. Therefore, canned pineapple manufactures must produce canned pineapple when the fresh pineapple fruits are abundant to secure a low raw material cost and top fruit quality.

There are variety of canned pineapple products depending upon the fruit colors ("choice" and "standard"), fruit cut (slice, chunk, and tidbit), can sizes, syrup sweetness level, and so forth. The choice color of pineapple refers to a deep dark yellow color of the pineapple meat. The color is preferred by most customers. Canned products made with the choice color pineapple are generally sold at a higher price than the same products made with the standard color fruit. However, the color of the pineapple cannot be identified until the fruit is peeled, but fresh pineapple is sold to the canned manufacturers in bulk and unpeeled. Only monthly ratios of pineapple with choice and standard colors can be estimated. Canned manufacturers must sometimes buy additional fresh pineapple to ensure that there are enough choice color fruits to serve the pre-ordered and future demands. Any leftover fruits after all demands are fulfilled must be processed right away as fresh pineapple is perishable by being canned and stored in the warehouse for future orders. Some leftover is discarded as waste because there is no room available in the warehouse or it is too costly to keep it as a safety stock. Under all these conditions, the objective of a canned pineapple manufacturer is to determine a production plan that maximizes the total profit.

Canned pineapple manufacturers are generally facing a production planning problem under seasonal supply of fresh pineapple. A number of decisions needs to be addressed in the planning, specifically multiple products (cuts and can sizes based on the available colors) to be manufactured over multiple planning periods and under a warehouse capacity constraint. There are also several production related costs involved, adding additional complexity to the problem.

Certain aspects of this production planning problem were studied by Kogan et al. (1996). Their planning was to be responsive to customer demands as much as possible with make-to-order production, while considering make-to-stock products and minimizing inventory and purchasing costs. Soman et al. (2006) tested a conceptual framework for production planning and inventory management in a food industry with a

combined make-to-stock and make-to-order production. Chen et al. (2014) examined pricing and production of a combined make-to-stock and make-to-order system. Any demands that could not be immediately satisfied were backlogged or lost. They focused on monotonicity of the optimal control policy and the optimal price. Grillo et al. (2017) formulated a model that aimed to maximize two conflicting objectives, total profit and mean product freshness, of a fruit supply chain.

This paper focus on multi-products multi-periods production planning for make-to-order demands in the canned pineapple industry in which the raw material—fresh pineapple—is perishable and available seasonally. A challenge in this research is to determine the amount of fresh pineapple to purchase while the colors of the pineapple vary each month.

The organization of this paper is as follows. The next section describes the problem in more details, and the mathematical model is formulated. Then a numerical example in which certain data are based on a canned pineapple manufacturer in Thailand is presented, and its results are discussed. Finally, the last section concludes the paper.

PROBLEM FORMULATION

Problem Description

Since pineapple is a seasonal fruit, its acquiring price is cheapest when it is in season. Like many other fresh fruits, pineapple is perishable and must be processed as soon as it is harvested, often by canning or drying. Canned pineapple has a larger market than the dried one because of its longer shelf life and industrial standards are more acceptable worldwide.

Canned pineapple products are influenced by at least three factors: the color of the fruit, the cut, and the size of the can. In addition to fresh pineapple cost, several other costs are involved in the production such as can cost, sugar (to make the syrup) cost, water cost, labor cost, energy cost, and inventory holding cost. Warehouse storage availability during multiple planning periods is also a common issue for any canned pineapple manufacturer. Certain demands of canned pineapple are pre-ordered several months ahead of the delivery date to secure the goods at reasonable prices. The production planner of the canned pineapple manufacturer must determine the quantity of the fresh pineapple to purchase as well as the types of canned pineapple products and their quantities to manufacture in each time period in order to maximize the total profits. It is possible that some products are not sold right away but are stored in the warehouse to serve future demands. The products are usually palletized when kept in the warehouse. Each pallet contains a different number of cans depending on the can size.

Mathematical Model

The following mathematical model is a system of equations established to reflect the multi-products multi-periods production planning problem described above. Its objective is to maximize the total profit.

Indices

i	Pineapple color	$i = 1,2,3,\ldots,I$
j	Pineapple cut	$j = 1,2,3,\ldots,J$
k	Can size	$k = 1,2,3,\ldots,K$
t	Month	$t = 1,2,3,\ldots,T$

Parameters

$Profit$	Total profit (baht)
$Revenue$	Total revenue (baht)
$Cost$	Total costs (baht)
D^t_{ijk}	Demand of canned pineapple with color i cut j can size k in month t (cans)
PO^t_{ijk}	Price of canned pineapple with color i cut j can size k in month t (baht)
$InvO^t_{ijk}$	On-hand inventory of canned pineapple with color i cut j can size k in month t (cans)
CC_k	Cost per can of can size k (baht)
CE_k	Energy cost to manufacture a can of pineapple in can size i (baht/can)
CL_{jk}	Labor cost to manufacture a can of pineapple with cut j in can size k (baht/can)
CP^t	Average cost of fresh pineapple per kilogram in month t (baht/kilogram)
CS_k	Sugar cost to manufacture a can of pineapple in can size k (baht/can)
CW_k	Water cost to manufacture a can of pineapple in can size k (baht/can)
H	Inventory holding cost per pallet per month (baht/pallet/month)
QC_k	Quantity of can size k per pallet (cans/pallet)
WP_k	Weight of fresh pineapple needed to fill a can of size k (kilogram)
δ	Proportion of fresh pineapple by weight that can be canned (percentage)
GP^t_i	Proportion of pineapple color i available in month t (percentage)
Capt	Maximum product quantity that can be manufactured in month t (cans)
QCO^t_{max}	Maximum inventory that can be kept in month t (pallets)

Decision Variables

X^t_{ijk}	Quantity of canned pineapple product with color i cut j can size k to be produced in month t (cans)
QP^t	Quantity of pineapple that is bought in month t (kilograms)

Objective Function

The objective function is to maximize the total profits. Some costs are included for the completion of the model and may not affect the decision. They are also used as a means to communicate within the case study manufacturer.

Maximize *Profit*

$$Profit = Revenue - Cost \qquad (1)$$

where *revenue* and *cost* are described by Equations (2) and (3)

$$Revenue = \sum_{t=1}^{T}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(PO^t_{ijk} \cdot D^t_{ijk}) \qquad (2)$$

$$Cost = \sum_{t=1}^{T}(CP^t \cdot QP^t) + \sum_{t=1}^{T}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(CC_k \cdot X^t_{ijk}) + \sum_{t=1}^{T}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(CE_k \cdot X^t_{ijk})$$
$$+ \sum_{t=1}^{T}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(CL_k \cdot X^t_{ijk}) + \sum_{t=1}^{T}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(CS_k \cdot X^t_{ijk}) + \sum_{t=1}^{T}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(CW_k \cdot X^t_{ijk})$$
$$+ \sum_{t=1}^{T}\sum_{k=1}^{K}\left(H \cdot \left\lceil \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} InvO^t_{ijk}}{QC_k} \right\rceil\right) \qquad (3)$$

where $InvO^t_{ijk}$ is described by Equations (4).

$$InvO^t_{ijk} = (InvO^{t-1}_{ijk} + X^t_{ijk}) - D^t_{ijk} \quad \forall i,j,k,t \qquad (4)$$

and let $InvO^0_{ijk} = 0$ for all i, j and k.

Constraints

1. For each color, the raw material usage cannot exceed the available raw material in each month:

$$\sum_{j=1}^{J}\sum_{k=1}^{K}(WP_k \cdot X^t_{ijk}) \leq \delta \cdot GP^t_i \cdot QP^t \quad \forall i,t \qquad (5)$$

2. The production cannot exceed the capacity in any month:

$$Cap^t \geq \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} X^t_{ijk} \qquad \forall t \qquad (6)$$

3. The stock kept must not exceed the warehouse space availability in each month:

$$\sum_{k=1}^{K}\left\lceil \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} InvO^t_{ijk}}{QC_k} \right\rceil \leq QCO^t_{max} \quad \forall t \qquad (7)$$

4. The monthly pre-ordered demands must be completely fullfilled:

$$InvO^{t-1}_{ijk} + X^t_{ijk} \geq D^t_{ijk} \quad \forall i,j,k,t \qquad (8)$$

5. The decision variables and parameters must satisfy the following conditions.

$$InvO^{t-1}_{ijk} \geq 0 \quad \forall i,j,k,t \qquad (9)$$
$$QP^t \geq 0 \qquad \forall t \qquad (10)$$
$$X^t_{ijk} \in integer \quad \forall i,j,k,t \qquad (11)$$
$$X^t_{ijk} \geq 0 \qquad \forall i,j,k,t \qquad (12)$$

NUMERICAL EXAMPLE

In this paper, only eight canned pineapple products are considered with the following variety: color $i = 1$ (choice), $i = 2$ (standard); cut $j = 1$ (slice), $j = 2$ (tidbit) and can size $k = 1$ (0.6 kilograms per can), $k = 2$ (3 kilograms per can). The planning covers four consecutive months ($t = 1,2,3,4$). The average prices of fresh pineapple in month 1 to 4 are 6.4, 6.5, 6.7, 7.0 baht per kilogram, respectively.

The can cost (CC), energy cost (CE), labor cost (CL), sugar cost (CS), and water cost (CW) of each product are shown in Table 1. A pallet can accommodate up to 20 smaller size cans ($k = 1$) or 10 of the larger ones ($k = 2$). Each pallet incurs about 40 baht per month as its inventory holding cost.

Table 1 : Production Costs by Product

Product			Production Cost (Bath/Can)				
i	j	k	CC	CE	CL	CS	CW
1	1	1	1	0.3	0.4	0.2	0.25
1	1	2	6	1.5	1.5	0.6	0.75
1	2	1	1	0.3	0.45	0.2	0.25
1	2	2	6	1.5	2	0.6	0.75
2	1	1	1	0.3	0.4	0.2	0.25
2	1	2	6	1.5	1.5	0.6	0.75
2	2	1	1	0.3	0.45	0.2	0.25
2	2	2	6	1.5	2	0.6	0.75

The price of different canned pineapple products (PO) often varies on a month basis and can be summarized in Table 2. The table also shows the monthly pre-ordered demands for all the canned products, and the ratio of choice-color pineapple estimated based on a monthly basis. This ratio is uncertain and may change year by year. Two scenarios of this ratio are explored. The first one is when the ratio varies in a larger range (LR) than from 0.3-0.7; and the other one is when it varies in a smaller range (SR) from 0.4-0.6.

Once a fresh pineapple is peeled and cored, only about 80% of the original weight is left to be processed and canned. Due to the warehouse space availability, the manufacturer may hold up to 4,500,000 cans in the warehouse in any time period, or an equivalence of 100,00 pallets.

Table 2: Monthly Price per Can and Pre-Ordered Demand by Product

Month	Product			PO (Baht/Can)	Demand (Can)	Ratio of Choice-color Pineapple	
	i	j	k			LR	SR
1	1	1	1	16	600,000	0.7	0.6
	1	1	2	25	-		
	1	2	1	16	640,000	0.7	0.6
	1	2	2	59	160,000		
	2	1	1	15	-	0.3	0.4
	2	1	2	24	200,000		
	2	2	1	16	320,000	0.3	0.4
	2	2	2	56	60,000		
2	1	1	1	20	860,000	0.6	0.5
	1	1	2	25	-		
	1	2	1	17	880,000	0.6	0.5
	1	2	2	59	240,000		
	2	1	1	15	-	0.4	0.5
	2	1	2	24	280,000		
	2	2	1	16	440,000	0.4	0.5
	2	2	2	51	80,000		
3	1	1	1	16	248,000	0.4	0.5
	1	1	2	25	328,000		
	1	2	1	17	244,000	0.4	0.5
	1	2	2	59	272,000		
	2	1	1	15	-	0.6	0.5
	2	1	2	24	320,000		
	2	2	1	17	480,000	0.6	0.5
	2	2	2	62	276,000		
4	1	1	1	15	1,000,000	0.3	0.4
	1	1	2	25	-		
	1	2	1	17	960,000	0.3	0.4
	1	2	2	59	180,000		
	2	1	1	15	-	0.7	0.6
	2	1	2	22	640,000		
	2	2	1	14	1,000,000	0.7	0.6
	2	2	2	62	-		

Results

The production planning problem above was solved using Excel Solver. The results are shown Table 3.

From the table, the results show that in both scenarios the production plan tends to take advantage of low average fresh pineapple costs in earlier months by over-manufacturing certain products in the months prior to the delivery date even though inventory holding costs are incurred. Let a triplet (i,j,k) represents a product with color i, cut j and can size k. In the SR scenario for example, the production of product (1,1,1) and product (2,1,2) in month 3 exceed their monthly demand. The excess quantities of these products together with additional production in month 4 are to serve their demands in month 4. Similarly, in the LR scenario the production of product (1,1,1) in month 3 can accommodate both demands for months 3 and 4. Another example is product (1,2,2) in the LR scenario. Its production is accumulated over months 1, 2, and 3 to serve the demand in months 3 and 4. The results of some other canned pineapple products in Table 3 also exhibit similar early production.

Table 3: Production Plan for Large and Small Range of Choice-Color Pineapplr Ratios Compared to Demands

Month	Product			Demand (Can)	Production Plan	
	i	j	k		LR	SR
1	1	1	1	600,000	600,000	965,599
	1	1	2	-	-	3,534
	1	2	1	640,000	640,000	640,002
	1	2	2	160,000	508,000	161,332
	2	1	1	-	-	-
	2	1	2	200,000	200,000	200,000
	2	2	1	320,000	320,000	320,000
	2	2	2	60,000	60,000	60,000
2	1	1	1	860,000	860,000	494,401
	1	1	2	-	-	7
	1	2	1	880,000	880,000	879,998
	1	2	2	240,000	324,000	238,668
	2	1	1	-	-	-
	2	1	2	280,000	280,000	280,000
	2	2	1	440,000	440,000	440,000
	2	2	2	80,000	80,000	145,554
3	1	1	1	248,000	448,125	1,248,000
	1	1	2	328,000	328,000	324,459
	1	2	1	244,000	244,000	244,000
	1	2	2	272,000	20,000	270,000
	2	1	1	-	-	-
	2	1	2	320,000	338,729	588,413
	2	2	1	480,000	480,000	480,000
	2	2	2	276,000	276,000	210,446
4	1	1	1	1,000,000	799,875	-
	1	1	2	-	-	-
	1	2	1	960,000	960,000	960,000
	1	2	2	180,000	-	180,000
	2	1	1	-	-	-
	2	1	2	640,000	621,271	371,587
	2	2	1	1,000,000	1,000,000	1,000,000
	2	2	2	-	-	-

From production planning results of both scenarios, the quantities of fresh pineapple to purchase each month must be determined to meet production plan as shown in Figure 1. From the figure, the quantities of fresh pineapple to purchase in both scenarios are the highest in month 3. This is because the price of the fresh pineapple increase in month 4, so it is worthwhile to manufacture some excess canned pineapple in month 3 and keep it in stock for a month before selling it in month 4. In general, the quantities of fresh pineapple to purchase in the SR scenario is lower or close to that in the LR one, except for month 3.
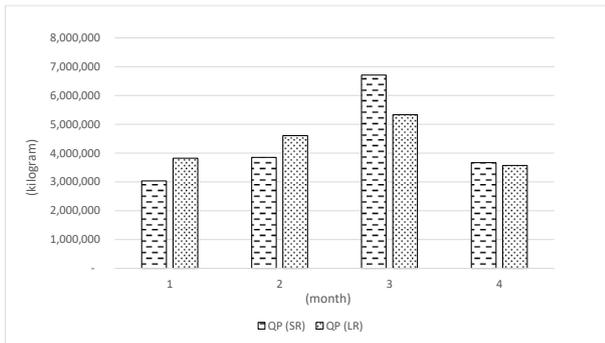
Figure 1: Quantity of fresh pineapple to be purchased each month in both scenarios (in kilogram)

Based on the pre-ordered demands and pre-determined revenue per can of all the canned pineapple products. The total revenue is 242,056,000 million baht. However, the production plans based on the two scenarios offer different total profit because they incur different costs, particular the pineapple cost and the inventory holding cost as shown in Table 4.

Table 4: Pineapple Cost, Inventory Holding Cost, and Profit in The Two Scenarios (In Baht)

Financial Item	Scenario	
	LR	SR
Pineapple Cost	114,571,382	114,449,393
Inventory Holding Cost	4,315,200	4,100,800
Other production Costs	48,799,600	48,799,600
Total Costs	167,686,183	167,349,793
Total Profit	74,369,817	74,706,207

Table 4 reveals that when the variation in the ratio of the choice-color pineapple fluctuates in a smaller range, the total profit is 336,390 baht higher than when this ratio varies in a larger range. This is because the fresh pineapple cost and the inventory holding cost in the SR scenario are lower than those in the LR scenario. A reason that the fresh pineapple (raw material) cost in the LR scenario is higher because in certain months the manufacturer needs to buy excess fresh pineapple so that there is enough choice-color pineapple to satisfy the product demands.

CONCLUSION

This paper studies a production planning problem that is normally experienced by a canned pineapple manufacturer. Pineapple is a seasonal fruit. Its price and availability of choice-color pineapple differ throughout the year. Several production and inventory holding costs are considered in this study. A mathematical model of this problem is formulated with eight different canned products under warehouse storage space limitation and pre-ordered demands over four consecutive months. The objective is to maximize the total profit. Two scenarios are experimented to examine how different ratios of

choice-color pineapple availability affect the planning and total profit.

The numerical results show that the production plans in both scenarios tend to over-manufacture some products when the price of the fresh pineapple is low and allow to incur inventory holding cost to keep the excess products until they are later needed to serve the demands. The ratio of choice-color pineapple availability slightly affects the total profit. The scenario when the ratio of choice-color pineapple availability varies within a smaller range offers a lower fresh pineapple cost and inventory holding cost than those of the scenario when this ratio fluctuates within a larger range.

This study focuses on make-to-order production. Due to seasonal availability of the fresh pineapple, a common practice in the canned pineapple industry is to make-to-stock some of the products to serve future demands. This is because the cost of the fresh pineapple which is the main ingredient of the canned pineapple products are the cheapest and its availability is peaked during the pineapple season. The optimal production plans obtained consider various production and inventory holding costs, estimated ratios of choice-color pineapple availability, and warehouse storage limitation to serve pro-ordered demands.

Excess fresh pineapple acquired so that there would be sufficient choice-color products to serve the customers could be used to produce make-to-stock products before it is wested. Future research may suggest make-to-stock quantities of different canned pineapple products based on the forecasted demands.

The results of this research can be used as a guide to canned pineapple production planning when there are changes in fresh pineaaple prices and color ratios. With this guide, a raw material purchasing plan and selling promotions could be set ahead of time.

REFERENCES

Chen, X., Tai, A.H. and Y. Yang. 2014. "Optimal production and pricing policies in a combined make-to-order/make-to-stock system," *International Journal of Production Research*, 52, 7027-45.

Grillo, H., Alemany, M.M.E., Ortiz, A. and V. S. Fuertes-Miquel. 2017. "Mathematical modelling of the order-promising process for fruit supply chains considering the perishability and subtypes of products," *Applied Mathematical Modelling*, 49, 255-78.

Kogan, K., Khmelnitsky, E. and O. Maimon. 1998. "Balancing facilities in aggregate production planning: Make-to-order and make-to-stock environments," *International Journal of Production Research*, 36, 2585-96.

Soman, C. A., van Donk, D.P. and G.J.C. Gaalman. 2007. "Capacitated planning and scheduling for combined make-to-order and make-to-stock production in the food industry: An illustrative case study," *International Journal of Production Economics*, 108, 191-99.

TRIDGE 2019 Overview of global value added pineapple market.https://www.tridge.com/intelligences/canned-pineapple/export (On-line accessed on 1 February 2021).

Wattanakul, T., Nonthapot, S., and T. Watchalaanun. 2020. "Factors determine Thailand's processed pineapple export competitiveness," *International Journal of Managerial Studies and Research*, 8(1), 36-41.

AUTHOR BIOGRAPHIES

KANAPATH PLANGSRISAKUL is a graduate student in Industrial and Manufacturing Systems Engineering program at the Department of Production Engineering, King Mongkut's University of Technology Thonburi, Thailand. His research interests are in production planning and optimization, and data management. His e-mail address is: `kanapath.002@mail.kmutt.ac.th`.

TUANJAI SOMBOONWIWAT is an Associate Professor in the Industrial Management section, Department of Production Engineering, Faculty of Engineering, King Mongkut's University of Technology Thonburi, Thailand. She received her M.Eng. in Industrial Engineering from Chulalongkorn University, Thailand, and Ph.D. in Industrial Engineering from Oregon State University, Corvallis, USA. Her research interests include green supply chain and logistics, business process and applications of operations research. Her e-mail address is: `tuanjai.som@kmutt.ac.th`.

CHAROENCHAI KHOMPATRAPORN holds a Ph.D. from University of Washington, USA. He is an Associate Professor in the Department of Production Engineering at King Mongkut's University of Technology Thonburi, Thailand. His research interests include supply chain and logistics management, applied operations research, optimization algorithms, and industrial sustainable operations management. He has published in several peer-reviewed journals, and worked closely with both public and private sectors such as hard disk drive manufacturer, automotive industry, and banking. His e-mail address is: `charoenchai.kho@kmutt.ac.th`

A ROBUST AND ADAPTIVE APPROACH TO CONTROL OF
A CONTINUOUS STIRRED TANK REACTOR WITH JACKET COOLING

Roman Prokop, Radek Matušů and Jiří Vojtěšek
Faculty of Applied Informatics
Tomas Bata University in Zlín
nám. T. G. Masaryka 5555, 760 01 Zlín, Czech Republic
E-mail: prokop@utb.cz

KEYWORDS

Robust Control, Adaptive Control, Continuous Stirred Tank Reactor, 2DOF Control Structure, Algebraic Control Design.

ABSTRACT

Continuous Stirred Tank Reactors (CSTR) are one of the main technological plants used in chemical and biochemical industry. These systems are quite complex with many nonlinearities and the conventional linear control with fixed parameters can be questionable or sometimes unacceptable. The solution should be found in so-called "non-traditional" control approaches like adaptive, robust, fuzzy, or artificial intelligent methods. One way is the utilization of self-tuning adaptive schemes, but computations may be quite difficult, clumsy and time-consuming. This paper brings an alternative principle called a robust approach and the comparison of the robust and adaptive control responses. Robust control considers a CSTR model as a linear system with parametric uncertainty, which covers a family of all feasible plants. Then several controllers with fix parameters are designed so that for all possible plants, the acceptable control behavior is obtained. The two-degree-of-freedom (2DOF) structure for the control law was chosen. Both robust and adaptive control is applied to an original nonlinear model of a CSTR. All calculations and simulations of mathematical models and control responses were performed in the Matlab and Simulink environment.

INTRODUCTION

The plants in technological processes and especially in chemical and biochemical industry usually have nonlinear behavior that causes difficulties in the control of such processes. Another unpleasant feature can be found in the complexity of such processes with a lot of variables and properties that result in difficult mathematical descriptions. This negative property should be overcome with the linearization of nonlinear models that introduces simplifications that reduces the intricacy of the system. On the other hand, this simplification can result in inaccurate descriptions of the system. The utilization of adaptive (e.g. self-tuning) schemes brings more difficult, clumsy and time-consuming computations (Åström and Wittenmark

1989). The control design using a hybrid adaptive control principle was used in (Vojtesek et al. 2017) where the originally nonlinear system was represented by an external linear model with recursively identified parameters and the pole-placement method adjustment principle was applied. A practically favored approach to overcome the loss of the model accuracy, compensated by its structure simplicity, consists in the utilization of a model with uncertainty. This idea allows working with the linear time-invariant low order mathematical models also for the case of real systems with complex dynamics or nonlinear behavior. There are several ways how to incorporate the uncertainty into the mathematical model available, see (Barmish 1994; Bhattacharyya et al. 1995). The popular group of uncertain systems is known as the systems with parametric uncertainty, which means the model structure is fixed but its parameters can vary, typically within some prescribed intervals. Then, the natural task is to find a controller, called a robust controller, that ensures the preserving some important closed-loop properties (e.g. stability) for the whole assumed family of controlled plants, see (Grimble 2006).

The system under the consideration is the Continuous Stirred Tank Reactor (CSTR) with the cooling in the jacket. The mathematical model of this system is described by the set of four nonlinear Ordinary Differential Equations (ODE). This set can be solved by standard numerical methods that are implemented in mathematical software such as Matlab, Simulink etc.

The main aim of this paper is in the design a robustly stabilizing controller for the CSTR with the cooling in the jacket, modelled as a system with parametric uncertainty, by means of algebraic approach. The work will put emphasis on the relatively easily tunable and applicable conventional PID controllers. The robust stabilization and control are verified and discussed by a simulation example of nonlinear CSTR.

CONTINUOUS STIRRED TANK REACTOR

The nonlinear controlled system under the consideration is a CSTR display of which can be found in Figure 1. The so-called Van der Vusse reaction described by general scheme:

$$A \xrightarrow{k_1} B \xrightarrow{k_2} C$$
$$2A \xrightarrow{k_3} D \qquad (1)$$
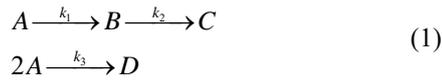
is performed inside the reactor.

This system can be described by a nonlinear mathematical model derived with the commonly used simplifications that reduce complexity of the system that has a lot of variables and connections. If we introduce these simplifications, the originally very complex system can be described by the set of nonlinear ordinary differential equations – see e.g. (Russell and Denn 1972) or (Vojtesek et al. 2017):

$$\frac{dc_A}{dt} = \frac{q_r}{V_r}(c_{A0} - c_A) - k_1 c_A - k_3 c_A{}^2$$

$$\frac{dc_B}{dt} = -\frac{q_r}{V_r} c_B + k_1 c_A - k_2 c_B$$

$$\frac{dT_r}{dt} = \frac{q_r}{V_r}(T_{r0} - T_r) - \frac{h_r}{\rho_r c_{pr}} + \frac{A_r U}{V_r \rho_r c_{pr}}(T_c - T_r) \qquad (2)$$

$$\frac{dT_c}{dt} = \frac{1}{m_c c_{pc}}(Q_c + A_r U(T_r - T_c)),$$

$$\text{where } 0 \leq c_A, \ 0 \leq c_B$$

This set is derivate with the help of material and heat balances inside the reactor. Variable t in the set of Ordinary Differential Equations (ODE) (2) denotes the time, c are concentrations, T represents temperatures, c_p is used for specific heat capacities, q_r means volumetric flow rate of the reactant, Q_c is heat removal of the cooling liquid, V are volumes, ρ stands for densities, A_r is the heat exchange surface and U is the heat transfer coefficient. Indexes $(\bullet)_A$ and $(\bullet)_B$ belong to compounds A and B, respectively, $(\bullet)_r$ denotes the reactant mixture, $(\bullet)_c$ cooling liquid and $(\bullet)_0$ are feed (inlet) values.
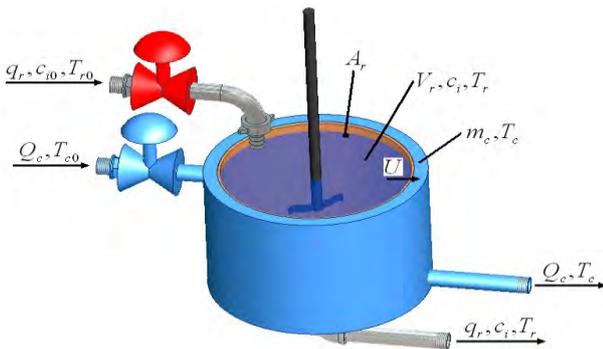
Figure 1: Continuous stirred tank reactor with cooling in the jacket

This reactor belongs to the class of *lumped-parameter nonlinear systems*, see e.g. Ingham et al. (2000). Nonlinearity can be found in reaction rates (k_j), which are described via the Arrhenius law:

$$k_j(T_r) = k_{0j} \cdot \exp\left(\frac{-E_j}{RT_r}\right), \text{ for } j = 1, 2, 3 \qquad (3)$$

where k_0 represent pre-exponential factors and E are activation energies.

The reaction heat (h_r) in Eq. (2) is expressed as:

$$h_r = h_1 \cdot k_1 \cdot c_A + h_2 \cdot k_2 \cdot c_B + h_3 \cdot k_3 \cdot c_A^2 \qquad (4)$$

where h_i means reaction enthalpies.

The initial conditions for the set of ODE (2) are

$$c_A(0) = c_A^s, \ c_B(0) = c_B^s, \ T_r(0) = T_r^s, \ T_c(0) = T_c^s \qquad (5)$$

The mathematical model of the system described by the set of ODE in Eq. (2) shows that this model has four state variables: $c_A(t)$, $c_B(t)$, $T_r(t)$ and $T_c(t)$. From the control point of view, several input variables can be used, e.g. input concentration of compound A, c_{A0}, input temperature of the reactant, T_{r0}, etc. However, the physical viability of these variables is greatly limited from the practical point of view. That is why are simulation studies mainly focused on the volumetric flow rate of the reactant q_r and the heat removal of the cooling liquid Q_c. The change of both quantities can be practically represented for example by the turn of the valve on the inlet pipe, or by the speed of the pump.

Fixed parameters of CSTR are given in Table 1.

Table 1: Parameters of CSTR

Name of the parameter	Symbol and value of the parameter
Volume of the reactor	$V_r = 0.01 \ m^3$
Density of the reactant	$\rho_r = 934.2 \ kg.m^{-3}$
Heat capacity of the reactant	$c_{pr} = 3.01 \ kJ.kg^{-1}.K^{-1}$
Weight of the coolant	$m_c = 5 \ kg$
Heat capacity of the coolant	$c_{pc} = 2.0 \ kJ.kg^{-1}.K^{-1}$
Surface of the cooling jacket	$A_r = 0.215 \ m^2$
Heat transfer coefficient	$U = 67.2 \ kJ.min^{-1}m^{-2}K^{-1}$
Pre-exponential factor for reaction 1	$k_{01} = 2.145 \cdot 10^{10} \ min^{-1}$
Pre-exponential factor for reaction 2	$k_{02} = 2.145 \cdot 10^{10} \ min^{-1}$
Pre-exponential factor for reaction 3	$k_{03} = 1.5072 \cdot 10^8 \ min^{-1}.kmol^{-1}$
Activation energy of reaction 1 to R	$E_1/R = 9758.3 \ K$
Activation energy of reaction 2 to R	$E_2/R = 9758.3 \ K$
Activation energy of reaction 3 to R	$E_3/R = 8560 \ K$
Enthalpy of reaction 1	$h_1 = -4200 \ kJ.kmol^{-1}$
Enthalpy of reaction 2	$h_2 = 11000 \ kJ.kmol^{-1}$
Enthalpy of reaction 3	$h_3 = 41850 \ kJ.kmol^{-1}$
Input concentration of compound A	$c_{A0} = 5.1 \ kmol.m^{-3}$
Input temperature of the reactant	$T_{r0} = 387.05 \ K$

STATIC AND DYNAMIC ANALYSES

Once we have mathematical model of the system, we can make simulation experiments that help with the understanding of the system's behaviour. Also, we can use this knowledge in the design of the controller which will be also described later in the Adaptive control section.

Steady-State Analysis

The steady-state analysis as the first step means that we want to know value of state variables, in our case concentrations c_A, c_B and temperatures T_r, T_c in so called steady-state. The mathematical meaning of this claim is the derivatives with respect to time in the set of ODE (2) are set to zero. It means that the set of ODE (2) is transformed to the set of nonlinear algebraic equations

$$c_A^s = \frac{-\left(\frac{q_r}{V_r} + k_1\right) \pm \sqrt{\left(\frac{q_r}{V_r} + k_1\right)^2 - \left(4 \cdot k_3 \cdot \left(-\frac{q_r}{V_r} c_{A0}\right)\right)}}{2 \cdot k_3};$$

$$c_B^s = \frac{k_1 \cdot c_A^s}{k_2 + \frac{q_r}{V_r}};$$

$$(6)$$

$$T_r^s = \frac{\frac{q_r}{V_r} T_{r0} - \frac{h_r}{\rho_r \cdot c_{pr}} + \frac{U \cdot A_r}{\rho_r \cdot c_{pr} \cdot V_r} T_c^s}{\frac{q_r}{V_r} + \frac{U \cdot A_r}{\rho_r \cdot c_{pr} \cdot V_r}};$$

$$T_c^s = \frac{Q_c}{U \cdot A_r} + T_r$$

That can be solved numerically for example with the use of simple iteration method. We can observe the steady-state behaviour for various input variables.

Results for various values of volumetric flow rate of the reactant, q_r, and heat removal of coolant, Q_c, are shown in Figure 2.
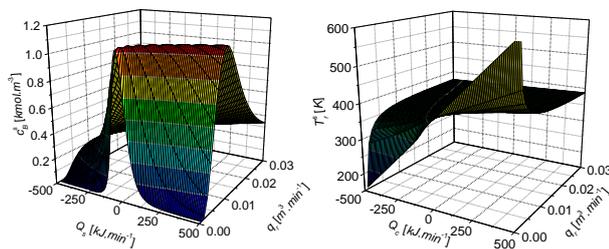
Figure 2: Steady-state analysis for various volumetric flow rate of the reactant, q_r, and heat removal of the cooling, Q_c

We can read from graphs, that this system has strongly nonlinear behaviour. The optimal working point can be represented by the combination of the volumetric flow rate of the reactant $q_r^s = 2.365 \cdot 10^{-3}$ $m^3.min^{-1}$ and the heat removal $Q_c^s = -18.56$ $kJ.min^{-1}$.

The dynamic analysis and the control is then performed around this working point where steady-state values of state variables are

$$c_A^s = 2.1403 \; kmol.m^{-3}, \quad c_B^s = 1.0903 \; kmol.m^{-3}$$
$$T_r^s = 387.34 \; K, \quad\quad T_c^s = 386.06 \; K$$

$$(7)$$

Dynamic Analysis

Once we have optimal working point from the steady-state analysis, we can continue with the dynamic analysis which means observing of the system's behaviour after the step change of the input variable. In our case, we have chosen the step changes of the coolant's heat removal, ΔQ_c, because this input will be than used as an action value for the control.

Investigated output variables are output concentration of the product B, $c_B(t)$, and output temperature of the coolant, $T_r(t)$. Both values are related to their steady-state values in (7) because we want to display these output from zero and as we can see in (5), these values are initial values in the numerical solution. Input and output variables are then:

$$u(t) = \frac{Q_c(t) - Q_c^s}{Q_c^s} \cdot 100 \; [\%]$$
$$y_1(t) = c_B(t) - c_B^s \; [kmol.m^{-3}] \quad\quad (8)$$
$$y_2(t) = T_r(t) - T_r^s \; [K]$$

Mathematically, the dynamic analysis means numerical solution of the set of ODE (2) together with (3) and (4). This numerical solution can be easily performed with build-in functions in Matlab or other mathematical software. Results are shown in Figure 3.
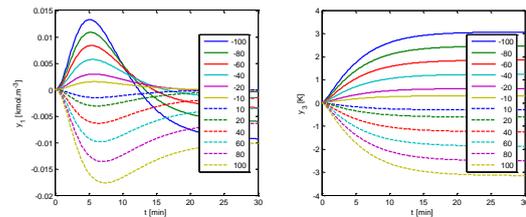

Figure 3: Results of the dynamic analysis for various step changes of input variable $u(t)$

Both courses of output variables $y_1(t)$ and $y_2(t)$ shows nonlinearity of the system which is obvious mainly for the output $y_1(t)$. On the other hand, output $y_1(t)$ can be expressed by the second order transfer function

$$G(s) = \frac{b(s)}{a(s)} = \frac{b_1 s + b_0}{a_2 s^2 + a_1 s + a_0} \quad\quad (9)$$

This output will be used as a controlled output in the control section of this paper.

ROBUST CONTROL

Models with Parametric Uncertainty

Systems with parametric uncertainty represent an effective and popular way of considering the uncertainty in the mathematical model of a real plant, see e.g. (Barmish 1994) or (Matušů and Prokop 2013; 2014). The utilization of such models supposes known structure (and order) of the transfer function but not precise knowledge of real parameters, which can be bounded by intervals with minimal and maximal possible values. They can be described by a transfer function:

$$G(s,q) = \frac{b(s,q)}{a(s,q)} \qquad (10)$$

where $b(s,q)$ and $a(s,q)$ denote polynomials in s (Laplace transform) with coefficients depending on q, which is a vector of real uncertain parameters. Typically, this vector is confined by some uncertainty bounding set, which is generally a ball in some appropriate norm. The combination of the uncertain system (e.g. transfer function (10)) with an uncertainty bounding set gives the so-called family of systems, see e.g. (Barmish 1994). A special and frequent case of a system with parametric uncertainty is an interval plant. Its parameters vary independently on each other within given bounds, i.e.:

$$G(s,b,a) = \frac{\sum_{i=0}^{m} \left[b_i^- ; b_i^+ \right] s^i}{\sum_{i=0}^{n} \left[a_i^- ; a_i^+ \right] s^i} \qquad (11)$$

where $b_i^-, b_i^+, a_i^-, a_i^+$ represent lower and upper limits for parameters of numerator and denominator, respectively.

Control Structure and Design

The 2DOF closed-loop control system with separated feedback and feedforward parts of the controller is depicted in Figure 4. The transfer functions $G(s)$, $C_b(s)$, and $C_f(s)$ represent controlled plant, feedback part of the controller, and feedforward part of the controller, respectively and the signals $w(s)$, $n(s)$, and $v(s)$ are reference, load disturbance, and disturbance signal.
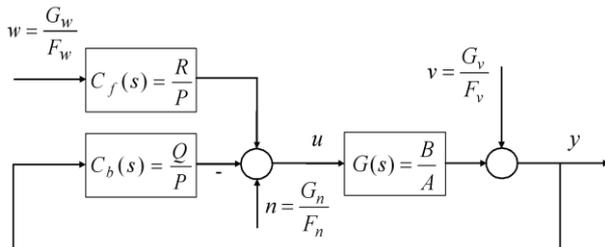
Figure 4: Two-degree-of-freedom control loop

The traditional (one degree of freedom) feedback system is obtained by $R=Q$. However, there are much relevant evidence that the feedforward part brings positive improvements in control responses, see e.g. (Gorez 2003) or (Matušů and Prokop 2013; 2014).

The control synthesis itself is based on the algebraic ideas of Vidyasagar (1985), and Kučera (1993). Subsequently, the specific tuning rules have been developed and analyzed e.g. in (Prokop and Corriou 1997) or (Matušů and Prokop 2013; 2014).

Besides, the controller tuning rules for the case of law order controlled plant under assumption of either purely reference tracking problem or reference tracking and load disturbance rejection together have been already studied e.g. (Kučera 1993) or (Matušů and Prokop 2013; 2014) and so this part presents the important results and then it is applied to the CSTR as a plant with parametric uncertainty.

First, the control design technique supposes the description of linear systems in Fig. 3 by means of the ring of proper and stable rational functions (R$_{PS}$). The conversion from the ring of polynomials to R$_{PS}$ can be performed very simply – see e.g. (Vidyasagar 1985) or (Prokop and Corriou 1997) according to:

$$G(s) = \frac{b(s)}{a(s)} = \frac{\dfrac{b(s)}{(s+m)^n}}{\dfrac{a(s)}{(s+m)^n}} = \frac{B(s)}{A(s)}, \qquad (12)$$

$$m > 0, \ n = \max\{\deg(a), \deg(b)\}$$

The parameter $m > 0$ will be later used as a controller-tuning knob. The value of the tuning knob has a relevant influence on the control behavior of control responses.

The algebraic analysis (Prokop and Corriou 1997; Matušů and Prokop 2013; 2014) leads to the first Diophantine equation:

$$A(s)P(s) + B(s)Q(s) = 1 \qquad (13)$$

with a general solution $P(s) = P_0(s) + B(s)T(s)$, $Q(s) = Q_0(s) - A(s)T(s)$, where $T(s)$ is an arbitrary member of (the ring) R$_{PS}$ and the pair $P_0(s)$, $Q_0(s)$ represents any particular solution of (13). Since the feedback part of the controller is responsible not only for stabilization but also for disturbance rejection, the convenient controller from the set of all stabilizing ones can be chosen on the basis of divisibility conditions. The requirement of the reference tracking is obtained by the second Diophantine equation (see Kučera, 1993, Matušů and Prokop, 2013, 2014):

$$F_w(s)Z(s) + B(s)R(s) = 1 \qquad (14)$$

Robust Stability

The stability of the feedback loop is a crucial requirement in all control applications. Naturally, the feedback loop can be stable when the controlled and/or control plant is unstable. In the case of uncertainty of controlled plants, robust stability means that not only one fixed closed-loop system is stable but also the whole

family of closed-loop control systems is ensured to be stable. Details can be found in e.g. (Ackermann 1993; Barmish 1994; Bhattacharyya et al. 1995; Matušů and Prokop 2011; 2013; 2014). This paper utilizes the robust stability tests based on a universal tool known as the value set concept in combination with the zero exclusion condition – see e.g. (Barmish 1994) or (Matušů and Prokop 2011).

ADAPTIVE CONTROL

The adaptive approach in this work is based on the recursive identification of the linearized model described by the transfer function (9) during the control. The control scheme is very similar to 2DOF control configuration in Figure 4 but block G is in this case mathematical model of the controlled system, in our case the set of ODE in (2).

The control synthesis employs pole-placement method together with the spectral factorization. Our previous experiments (for example (Vojtěšek and Dostál 2005; 2016)) have shown, that this method produces sufficient control results.

This control synthesis is based on the solution of the set of Diophantine equations

$$a(s)f(s)_i^\sim \qquad)q(s) = d(s)$$
$$t(s)f_w(s) + b(s)r(s) = d(s) \qquad (15)$$

where polynomials $a(s)$ and $b(s)$ are polynomials from the transfer function (9) and they are estimated recursively with the Ordinary recursive least-squares method (Bobál et al. 2005). Polynomial $t(s)$ is auxiliary polynomial and unknown controller's polynomials $p(s)$, $q(s)$ and $r(s)$ are computed from (16).

Unknown stable polynomial $d(s)$ on the right side of equations (16) was designed with the use of pole-placement method, e.g. this polynomial is generally

$$d(s) = \prod_{i=1}^{\deg d(s)} (s + s_i) \qquad (16)$$

where $s_i = \alpha_i + \omega_i j$ are roots of the polynomial and choice of these roots affects control results. More details about this method can be found for example in (Vojtěšek and Dostál 2005).

SIMULATIONS AND DISCUSSION

A Robust Approach

The CSTR was identified in (Vojtěšek et al. 2017) as a second order system with the transfer function (9) with nominal parameters: $a_2 = 1$, $a_1 = 1.4550$, $a_0 = 0.3072$, $b_1 = -0.0037$, $b_0 = -0.0095$. The intervals for uncertain perturbations were obtained by deeper analysis of the dynamic behavior and they result in the following ones:

$$a_0 = [0.24576;\ 0.36864],$$
$$a_2 = [0.8;\ 1.2], a_1 = [1.164;\ 1.746], \qquad (17)$$
$$b_1 = [-0.00296;\ -0.00444], b_0 = [-0.0076;\ -0.0114]$$

Three 2DOF controllers have been designed for the nominal plant and the tuning parameters. The firet one was generated for $m = 0.5$, the second one for $m = 0.8$ and the third one for $m = 1.2$. The feedback and feedforward parts of the controller for the first one is:

$$C_b(s) = \frac{q_2 s^2 + q_1 s + q_0}{s^2 + p_1 s} = \frac{-61.3653 s^2 - 39.7878 s - 6.5789}{s^2 + 0.3179 s}$$
$$C_f(s) = \frac{r_2 s^2 + r_1 s + r_0}{s^2 + p_1 s} = \frac{-26.3158 s^2 - 26.3158 s - 6.5789}{s^2 + 0.3179 s} \qquad (18)$$

Figure 5 and Figure 6 show the controlled and control variables for all three tuning parameters. The red lines depict the nominal plant responses and black shadows are responses for the whole uncertain family (17), represented by $3^5 = 243$ members (three values for each interval parameter: minimum, midpoint, and maximum). The load disturbance $n = 10$ was injected in the time $t = 150$ and it is evident that no permanent error is observed.
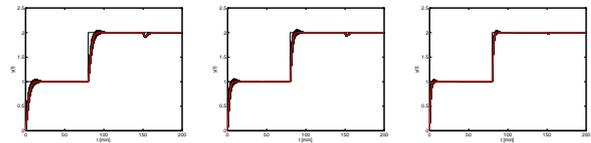

Figure 5: Set of output controlled variables for m=0.5 (left), m=0.8 (middle), and m=1.2 (right)
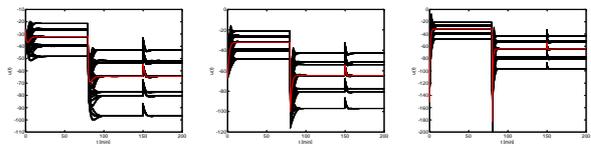

Figure 6: Set of input control variables for m=0.5 (left), m=0.8 (middle), and m=1.2 (right)

Simulation results proved that the fix robust controller could be designed for a wide family of interval systems. The results are shown in Figures 5 and 6 for three values of the tuning parameter m>0. The choice of the tuning parameter m>0 was found empirically and experimentally. Until now, there is no exact theory on how to obtain the optimal value (see e.g. Prokop and Corriou, 1997). The Figure 7 shows the zoomed value sets for all three values of m. All three subfigures from Figure 7 may seem the same for the first sight, but please note the differences in axes ranges. Anyway, they confirm the robust stability of the designed control loops since they are excluded from the critical point $(0,0j)$ and all required preconditions are fulfilled (Barmish 1994).
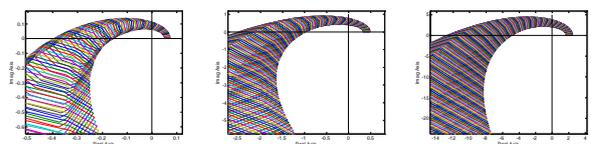

Figure 7: Zoomed value sets for m=0.5 (left), m=0.8 (middle) and m=1.2 (right)

In order to verify the practical usability of the designed controllers, they were applied not only to the linearized model, but also to the original nonlinear model of CSTR. The control results for this nonlinear case are shown and mutually compared in Figure 8.
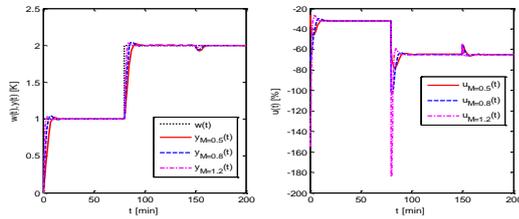

Figure 8: Robust control of the original nonlinear model for three values of *m* – comparison of the output controlled variables (left) and the input control variables (right)

The control results shown in Figure 8 assumes that there is no limitation of the control signal. On the other hand, Figure 9 provides the control behavior for the same controllers, but with the saturated control signals in the range ±100 %. It can be seen that this saturation affects the signals for *m*=0.8 and *m*=1.2. Consequently, higher peaks caused by the wind-up effect are observable, especially for *m*=1.2.
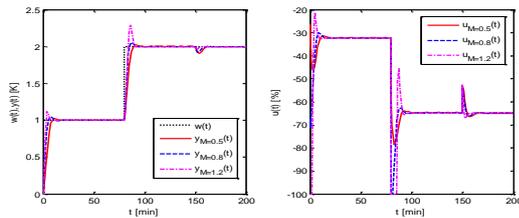
Figure 9: Robust control of the original nonlinear model for three values of *m* and the saturated control signal (±100 %) – comparison of the output controlled variables (left) and the input control variables (right)

An Adaptive Approach

Three adaptive controllers for 2DOF configuration were tuned, assuming the placement of the closed-loop poles 0.07, 0.1, and 0.2, respectively. Figure 10 shows the control results for the original nonlinear CSTR model, and Figure 11 presents the evolution of the identified parameters during the simulation.
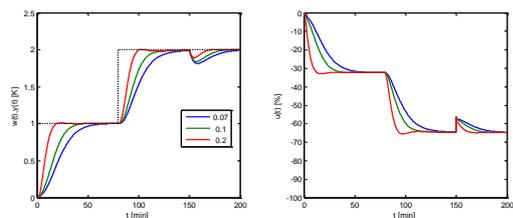
Figure 10: Adaptive control results – output (left) and control (right) signals
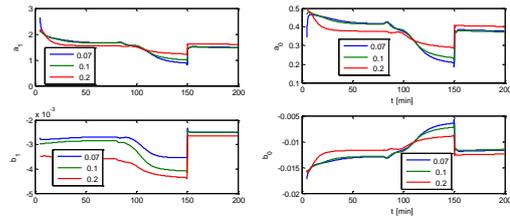
Figure 11: Adaptive control results – identified parameters

Comparison and Discussion

The control performance of both robust and adaptive approaches can be tuned by the parameter *m* or the proper pole-placement. In all cases, the costs for the rapid control and better disturbance rejection are the higher and more aggressive control signals. For some faster robust controllers, the control signals would have to be restricted for the practical application. The main advantage of the self-tuning controllers is obvious from its name, i.e., after successful initialization, they are able to control the CSTR without knowledge of the model. On the other hand, the main advantage of the off-line tuned robust controllers is their simplicity and reliability, even under prescribed model uncertainty. The comparison of control results for two reasonable choices of tuning parameters, i.e., *m*=0.5 for the robust controller, and α=0.2 for the adaptive controller, are shown in Figure 12. Anyway, it was shown that both approaches are able to control the CSTR satisfactorily and the final choice of the approach depends on the additional requirements or preferences of a user or control engineer.
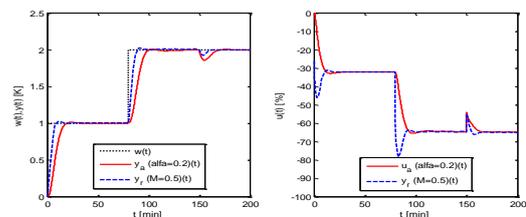

Figure 12: Comparison of selected robust and adaptive controllers – output (left) and control (right) signals

CONCLUSIONS

Modelling and control of CSTR are addressed in the contribution. Two different approaches of control were designed and compared. The first one is an adaptive self-tuning principle based on the recursive identification procedure with polynomial control design. The second control principle utilizes robust control algorithms designed in the ring R_PS. The synthesis method itself is based on linearized model with parametric uncertainty and accompanied by the analysis of robust stability. Both approaches use the 2DOF feedback control structure. As an application, a set of designed robust and adaptive controllers were applied to

control of an original nonlinear model of CSTR. The main aim of the control design was energy saving in the industry operation of CSTR. All simulations were performed in the MATLAB and Simulink environment.

REFERENCES

Ackermann, J. 1993. *Robust control – systems with uncertain physical parameters*. Springer-Verlag London, Great Britain.

Åström, K.J., Wittenmark, B., 1989. Adaptive Control. Addison Wesley. Reading. MA, USA. ISBN: 0-201-09720-6.

Barmish, B. R., 1994. *New Tools for Robustness of Linear Systems*, New York, USA: Macmillan.

Bhattacharyya, S.P., Chapellat, H., Keel, L.H., 1995. Robust Control-The Parametric Approach. Prentice Hall.

Bobál, V., Böhm, J., Fessl, J., Macháček, J. 2005. Digital Self-tuning Controllers. Algorithms, Implementation and Applications. Springer 2005.

Gorez, R., 2003. New design relations for 2-DOF PID-like control systems. *Automatica*, vol. 39, no. 5, pp. 901-908,

Grimble, M. J., 2006. Robust Industrial Control Systems. Optimal Design Approach for Polynomial Systems. Prentice Hall. USA. ISBN: 0-470-02073-3

Ingham, J., Dunn, I. J., Heinzle, E., Prenosil, J. E., 2000. Chemical Engineering Dynamics. An Introduction to Modeling and Computer Simulation. Second. Completely Revised Edition. VCH Verlagsgesellshaft. Weinheim, Germany. ISBN: 3-527-29776-6

Kučera, V., 1993. Diophantine Equations in Control – A survey. Automatica. 29, 1993, 1361-1375.

Matušů, R., Prokop, R., 2011. Graphical analysis of robust stability for systems with parametric uncertainty: an overview. *Transactions of the Institute of Measurement and Control*, Vol. 33, No. 2, pp. 274-290.

Matušů, R., Prokop, R., 2013. Algebraic Design of Controllers for Two-Degree-of-Freedom Control Structure. *International Journal of Mathematical Models and Methods in Applied Sciences*, vol. 7, no. 6, pp. 630-637.

Matušů, R., Prokop, R., 2014: An Algebraic Approach to Two-Degree-of-Freedom Controller Design for Systems with Parametric Uncertainty. *International Journal of Mathematical Models and Methods in Applied Sciences*, Vol. 8, pp. 113-120, ISSN 1998-0140.

Middleton, H., Goodwin, G. C., 2004. Digital Control and Estimation - A Unified Approach. Prentice Hall. Englewood Cliffs, USA. ISBN: 0-13-211798-3

Prokop, R., Matušů, R., Vojtěšek, J., 2019. Robust Control of Continuous Stirred Tank Reactor with Jacket Cooling. *Chemical Engineering Transactions*, Vol. 76, 2019, pp. 787-792, ISSN 2283-9216.

Russell, T., Denn, M. M., 1972. Introduction to Chemical Engineering Analysis. New York: Wiley, USA, xviii, ISBN: 04-717-4545-6.

Vidyasagar, M., 1985. Control system synthesis: A factorization approach, Cambridge, Massachusetts, USA: MIT Press.

Vojtěšek, J., Dostál, P., 2016. Continuous-time vs. discrete-time identification models used for adaptive control of nonlinear process. In: Proceedings - 30th European Conference on Modelling and Simulation, ECMS 2016 [online]. European Council for Modelling and Simulation (ECMS), 2016, s. 320-326.

Vojtěšek, J., Prokop, R., Dostál, P., 2017. Two Degrees-of-Freedom Hybrid Adaptive Approach with Pole-placement Method Used for Control of Isothermal Chemical Reactor. *Chemical Engineering Transactions*, 2017, Vol. 2017, No. 61, pp. p1-p7. ISSN 2283-9216.

AUTHOR BIOGRAPHIES

ROMAN PROKOP was born in Hodonín, Czech Republic in 1952. He graduated in Cybernetics from the Czech Technical University in Prague in 1976. He received post graduate diploma in 1983 from the Slovak Technical University. Since 1995 he has been at Tomas Bata University in Zlín, where he presently holds the position of full professor of the Department of Automation and Control Engineering and a vice-dean of the Faculty of Applied Informatics. His research activities include algebraic methods in control theory, robust and adaptive control, autotuning and optimization techniques. His e-mail address is: `prokop@utb.cz`.

RADEK MATUŠŮ received the M.S. degree the Faculty of Technology, Tomas Bata University (TBU) in Zlín, in 2002, and the Ph.D. degree from the Faculty of Applied Informatics (FAI), TBU in Zlín, in 2007. He was appointed an Associate Professor of machine and process control at FAI TBU in Zlín, in 2018. He has been holding various research or pedagogical positions at TBU in Zlín, since 2004, where he is currently a Researcher and a Project Manager. His research interests include analysis and synthesis of robust control systems, fractional-order systems, and algebraic methods in control design. He has (co-)authored more than 50 scientific journal papers and over 110 conference contributions. He serves as an Academic Editor or a Reviewer for dozens of scientific journals.

JIRI VOJTESEK was born in Zlin, Czech Republic. He studied at Tomas Bata University in Zlin, Czech Republic, where he received his M.Sc. degree in Automation and control in 2002. In 2007 he obtained Ph.D. degree in Technical cybernetics at Tomas Bata University in Zlin. In the year 2015 he became an associate professor. His research interests are modeling and simulation of continuous-time chemical processes, polynomial methods, optimal, adaptive and nonlinear control. You can contact him on e-mail address vojtesek@utb.cz.

CAPACITY LOSS ESTIMATION FOR LI-ION BATTERIES BASED ON A SEMI-EMPIRICAL MODEL

*Mohammed Rabah, Eero Immonen, and Sajad Shahsavari
Computational Engineering and Analysis (COMEA)
Turku University of Applied Sciences
20520 Turku, Finland
Email: *mohamed.rabah@turkuamk.fi

Mohammad-Hashem Haghbayan
Department of Future Technologies
University of Turku (UTU)
20500 Turku, Finland

Kirill Murashko
Department of Environmental and Biologic Science
University of Eastern Finland
70211 Kuopio, Finland

Paula Immonen
Laboratory of Electrical Engineering
LUT University
53850 Lappeenranta, Finland

Keywords

SEM (Semi-Empirical Model), LIBs (Li-ion batteries), C_{loss} (Capacity Loss), Cycling Aging, Calendar Aging.

ABSTRACT

Understanding battery capacity degradation is instrumental for designing modern electric vehicles. In this paper, a Semi-Empirical Model for predicting the Capacity Loss of Lithium-ion batteries during Cycling and Calendar Aging is developed. In order to predict the Capacity Loss with a high accuracy, battery operation data from different test conditions and different Lithium-ion batteries chemistries were obtained from literature for parameter optimization (fitting). The obtained models were then compared to experimental data for validation. Our results show that the average error between the estimated Capacity Loss and measured Capacity Loss is less than 1.5% during Cycling Aging, and less than 2% during Calendar Aging. An electric mining dumper, with simulated duty cycle data, is considered as an application example.

INTRODUCTION

The transport sector is one of the largest global emitters of carbon dioxide (CO_2), accounting for about 22% of the total emission (Kluschke et al., 2019). Electric cars have proven to be an efficient way of reducing these emissions in *passenger transport*, but electrification of *heavy-duty vehicles* (e.g. trucks, forest harvesters and mining dumpers) is more challenging, mainly due to limitations in battery technology. Among others, Liimatainen et al. (2019) concluded that battery electric trucks have not been a viable option to replace traditional diesel-powered ones because of the high energy requirements and low energy density of batteries. Furthermore, the absence of charging facilities in off-road conditions may render electrification of forest harvesters impractical. On the other hand, heavy-duty vehicles are typically tailored for a specific application niche, and their production batches are much smaller than those for passenger cars, which means that application-specific

design optimization is both necessary and can also have a significant effect on the vehicle performance — and thus on business profitability. At the heart of this design optimization is an understanding of the performance of lithium-ion (Li-ion) batteries.

One major difference between internal combustion engine vehicles (ICEV) and battery electric vehicles (BEV) is that the energy system in the latter degrades during use. While the performance of a diesel engine remains largely unaffected by repeated refuels and use, this is not the case for Li-ion batteries (LIBs): The capacity of LIBs decreases in both repeated cycling and storage. Moreover, the LIB degradation process depends on the battery chemistry and the way (or path) of usage. Dubarry and Devie (2018) concluded that the LIB cell temperature history had the strongest impact on degradation followed by the C-rate (i.e. charge/discharge current) and the state of charge (SoC). Also, they found that LIBs lose capacity faster at low SoCs during calendar aging and under small SoC swings while under cycling.

It is obvious from the above discussion that design optimization for heavy-duty battery electric vehicles (HDBEV) must address battery degradation. The upshot is that, in contrast to passenger cars, since a HDBEV is designed for a specific application, the typical usage conditions — including temperature, SoC and C-rate — can often be estimated with more accuracy during design. Battery system design optimization for HDBEVs thus requires parametric mathematical models of battery aging, estimated from real-world cycling and storage tests. The purpose of this article is to address this concern.

In this article, a Semi-Empirical Model (SEM) is proposed for estimating the capacity loss (C_{loss}) for different LIB chemistries during cycling and calendar aging. The model is developed based on the effect of four different parameters, namely temperature, time, depth of discharge, and C-rate current. The model is able to estimate the C_{loss} of LIBs with a high accuracy and low computation complexity compared to the other models. This model can be used for optimizing LIB systems for different chemistries throughout their lifetime
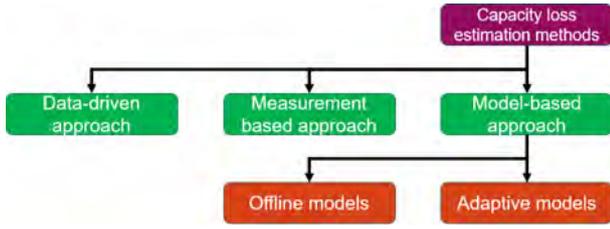
Figure 1: Capacity Loss Estimation methods.

in realistic operation conditions. The contribution of this article is elaborated in the next section.

RELATED WORK

LIB degradation is usually measured by C_{loss}. It reflects the ability of the LIB to store and to supply energy relative to its initial conditions, considering the energy and power requirements of the application (Berecibar et al., 2016). Many different methods for the C_{loss} estimation were presented in recent years and in general they can be divided into three approaches, which are shown in Figure 1.

The measurement based methods by themselves cannot be used for the estimation of the C_{loss} during design of the battery system as these methods requires the analysis of the measured data during operation of the LIB.

The data-driven approach includes methods, where the previously collected information about operation of the LIB is used to find the dependency between the collected information and C_{loss}. Different algorithms such as Support Vector Regression algorithm (Liu et al., 2020), Fuzzy logic (Yang et al., 2020) and Neural Networks (Naha et al., 2020) were investigated for C_{loss} estimation. These methods allows to estimate the C_{loss} with good accuracy in case of the availability of a sufficient amount of previously collected data. However, these methods are computationally intensive and are sensitive to the size and quality of the experimental dataset applied during the training, which are not always available.

The model-based approach (Cacciato et al., 2016) may yield better results for the C_{loss} estimation in case of the low amount of the experimental data. It can be done by applying a model, which should describe the processes occurring in the LIBs. Such models can be used offline and they directly provide the required information on the C_{loss} — or the required information can be obtained during comparison of the calculated and measured data in real time. If a very high accuracy for the C_{loss} estimate is needed, the second type of the battery model, so called adaptive model (Cen & Kubiak, 2020), is often recommended in the literature. The adjustment of the adaptive model parameters and C_{loss} estimation after comparison with measured data can be done by using such algorithms as Kalman Filter, Particle filter, Sliding mode observer etc. As it was reported by (Andre et al., 2013), the use of the adaptive models allows to estimate C_{loss} and State of Charge (SoC) simultaneously with estimation error under 1%.

Despite the high accuracy of the C_{loss} estimation, a high computational complexity may limit the applicability of the adaptive models during battery system design where high accuracy is not always necessary. In this case, offline models such as SEM (Singh et al., 2019) may be more useful as they may estimate the C_{loss} with acceptable uncertainty in case of the lack of the experimental data and they have low computation complexity. In the SEM approach, one attempts to identify a (simple) parametric function that describes the capacity reduction, through parametric optimization.

The applicability of the SEM approach for the C_{loss} estimation during storage (Grolleau et al., 2014) and cycling (Bocca et al., 2015) were widely shown for different LIBs. However, the presented models were usually verified for the same LIBs, from which the SEMs were created and the use of the presented algorithms for the creation of the SEMs for other type of the LIBs is not well discussed. Therefore, the research work described in the present article focused on the analysis of the applicability of the commonly used approach for the creation of the SEMs of different LIBs at different operation conditions.

SEM SPECIFICATIONS

In this article the most commonly used models, which can estimate the capacity loss C_{loss} of the LIBs during Calendar Aging C_{loss}^{cal} and Cycling Aging C_{loss}^{cyc} are analysed. These semi-empirical models were previously used for the modeling of the C_{loss} in LFP cells (Wang et al., 2011), NMC cells (Schmalstieg, Käbitz, Ecker, & Sauer, 2014), NCA cells (Petit, Prada, & Sauvant-Moynot, 2016) and they are briefly described below.

Calendar Aging

The two main factors that affects the Calendar Aging are the T and SoC. The general equation for the Calendar Aging estimation can be presented as:

$$C_{loss}^{cal} = B(SoC) \cdot e^{-\frac{E}{R(T-T_{ref})}} \cdot t^z, \qquad (1)$$

where B is the pre-exponential factor that depends on SoC, T is the temperature expressed in K, T_{ref} is the reference temperature also expressed in K and is equal to 298.15, R is the gas constant, E is the activation energy of a reaction, expressed in J/mol, t is the time in days, and z is a constant. The pre-exponential factor B can be presented as:

$$B = a_1 \cdot SoC + a_2 \qquad (2)$$

Where a_1 and a_2 are fitting constants. Equation 1 can be used to estimate the C_{loss} of the LIB during long period storage.

Cycling Aging

For Cycling Aging, the C_{loss} is mainly affected by current I, T and number of cycles N. Furthermore, other parameters do have a margin effect depending on the temperature of the LIB, e.g. depth of discharge DoD,

and the rated capacity. The general equation used to estimate the C_{loss}^{cyc} is as follow:

$$C_{loss}^{cyc} = B_{cyc}(I) \cdot e^{-\frac{E+\alpha \cdot |I|}{R(T-T_{ref})}} \cdot A_h^{z_{cyc}} \qquad (3)$$

Where B_{cyc} is a pre-exponential factor which depends on cycling current I, α and z_{cyc} are the fitting coefficients, and A_h is the full used capacity that can be obtained using the following equation:

$$A_h = FCE \cdot C_r = N \cdot DoD \cdot C_r \qquad (4)$$

Where FCE is the full cycle equivalent, C_r is the rated capacity.

MODEL IDENTIFICATION

The process of model identification is divided into two parts; (a) Data Selection and Fitting, and (b) Model Validation. These two parts are illustrated in Figure 2.
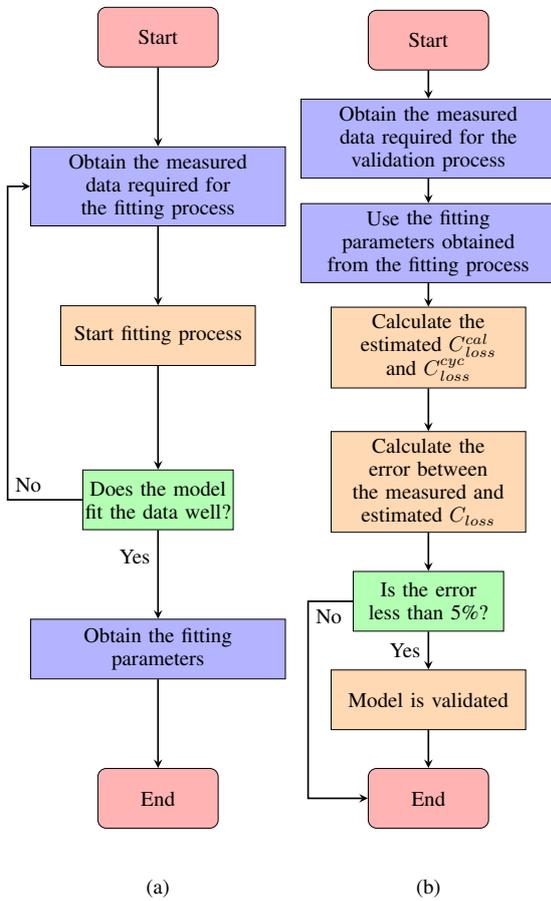
Figure 2: SEM flowchart process; (a) Data Selection and Fitting process, (b) Model Validation process.

Data Selection and Fitting

Figure 2 (a) shows the Flow Chart of the data selection and fitting process. As shown in this Figure, measured data from different references, specifically the C_{loss}^{cal}, C_{loss}^{cyc}, FCE and $Time$ are obtained. Afterwards the fitting process is applied to estimate the value of the fitting parameters. The algorithm used for obtaining the fitting parameters is shown in Algorithm 1. In this algorithm, $Mes\_C_{loss}^{cyc}$ and $Mes\_C_{loss}^{cal}$ are

---

Algorithm 1 Data Selection and Fitting process

Input: $Mes\_C_{loss}^{cyc}$, $Mes\_C_{loss}^{cal}$
Output: $Est\_C_{loss}^{cyc}$, $Est\_C_{loss}^{cal}$, R,a_1,a_2,z,α,z_{cyc},B_{cyc}

Body:

1: **measured** ← {**Mes_C$_{loss}^{cyc}$**, **Mes_C$_{loss}^{cal}$**}
2: **estimated** ← {**Est_C$_{loss}^{cyc}$**, **Est_C$_{loss}^{cal}$**}
3: **for** i ← 1 **to** 1000 **do**
  estimated ← $fitting$(**measured**)
  if estimated == **localminimum then**
    | Stop
  end
end
4: **error** ← $avg.(\frac{\text{estimated}-\text{measured}}{\text{measured}})$
5: **if error** < 0.05 **then**
  | {**E**, **R**, **a$_1$**, **a$_2$**, **z**, α, **z$_{cyc}$**, **B$_{cyc}$**} ← $fitpar$(**estimated**)
end

---

the measured C_{loss} data needed for fitting the model, while $Est\_C_{loss}^{cyc}$ and $Est\_C_{loss}^{cal}$ are the estimated C_{loss} output. In this process, the `fitting` function is using the `fminsearch` from MATLAB which uses a simplex search method (Lagarias et al., 1998) to obtain the estimated data output. In order to decrease the overall error between the measured and the estimated output, the fitting function is proceed in different scenarios, e.g. during a fixed temperature, fixed $DoD = 1 - SoC$, etc., for 1000 iteration in each. For this, several fitting parameters are calculated according to each situation. Once the estimated output reaches the local minimum, the average error is calculated between the estimated output and the measured data. In this case, it's assumed that 5% is when the model does fit the data well. Afterwards, function `fitpar` which uses the `polyfit` from MATLAB is used to generate the fitting parameters based on least square regression. Once these values are calculated, the model validation process is started as shown in Figure 2 (b).

Algorithm 2 Model Validation process

Input: $NewMeasuredData$
Output: C_{loss}^{cyc}, C_{loss}^{cal}
Parameters: R,a_1,a_2,z,α,z_{cyc},B_{cyc},T_{ref}

Body:

1: {**SoC**, **T**, **I**, **t**, **N**, **period**, **C$_r$**} ← **NewMeasuredData**
  // Calculate C_{loss} for Cycling Aging
2: **cyc.DoD**, **cyc.SoC**, **ch1**, **ch2** ← rainflow(**SoC**);
  cyc.I ← $avg.$ **I** $from$ **ch1** to **ch2**;
  cyc.T ← $avg.$ **T** $from$ **ch1** to **ch2**;
3: **C$_{loss}^{cyc}$** ← $B_{cyc} \cdot e^{-\frac{E+\alpha \cdot |cyc.I|}{R(cyc.T-T_{ref})}} \cdot (NC_r(cyc.DoD))^{z_{cyc}}$
  // Calculate C_{loss} for Calendar Aging
4: **if I** == 0 **then**
  cal.SoC = $avg.$ (**SoC**);
  cal.T = $avg.$ (**T**);
  cal.t = **period** * 24 * 3600;
  B(SoC) = **a$_1$** · **cal.SoC** + **a$_2$** ;
  C$_{loss}^{cal}$ ← $B(SoC) \cdot e^{-\frac{E}{R(cal.T-T_{ref})}} \cdot cal.t^z$
end

Model Validation

In this process, the estimated fitting parameters values are input into Equations (1)-(4), to calculate the C_{loss} during Cycling/Calendar aging, and the output

is compared to a known reference, where the error is calculated between the known C_{loss} and the estimated one. If the error is low (here defined as below 5%), this shows that the model is validated and can be used to estimate the C_{loss} of an experimental LIB.

The method used for calculating the C_{loss} is shown in Algorithm 2. The information about state of charge SoC, temperature T, current I, rated capacity C_r and time between operation cycles is necessary for the C_{loss} calculation t. The loss of the capacity during cycling is calculated from information about N, cycle start time $ch1$ and end time $ch2$, average values of $cyc.DoD$ and $cyc.SoC$ that are calculated from SoC curve by using `Rainflow` algorithm in MATLAB (ASTM E1049-85, 2005). Afterwards, average values of temperature $cyc.T$ and current $cyc.I.$ are calculated during a cycle to be used in Equation 3. The calendar C_{loss} is calculated by considering the average state of charge $cal.SoC$ and average temperature $cal.T$ of each long enough period of time when LIBs are not used where there is no current usage during this period.

RESULTS

To test the feasibility of the proposed model, several LIB chemistries should be evaluated. In this work, two different chemistries of LIBs have been chosen; Lithium Iron Phosphate (LFP) and Lithium-Titanate Oxide (LTO). These chemistries are among the primary candidates for modern HDBEV systems.

Lithium-titanate battery (LTO)

Data Selection and Fitting

The measured data that is used for the fitting process is acquired from (Dubarry & Devie, 2018). In his work, he studied the effect of temperature 'T', SoC swing range '$\triangle SoC$', and C-rate 'C' on the battery cells to measure its C_{loss} during 1400 (1C rate) to 4200 (3C rate) full cycle equivalent as demonstrated in Figure 3(a). Likewise, he studied the effect of 'T' and SoC through 61 weeks of Calendar Aging as illustrated in Figure 3(b).

Once the measured data from Figure 6 is extracted, the fitting process is started. The fitting function for the Cycling Aging is proceed in four different scenarios: during a fixed $DoD(40\%)$, fixed temperature (25°C), and fixed C-rate. Figure 4 shows the estimated C_{loss} (dashed line) compared to the measured one (solid line). The average error between the estimated C_{loss} and measured C_{loss} is found to be 0.63% at 50% of FCE and 0.54 at 100% of FCE, except for one point that shows an error of 0.72% during the 4200 FCE that can be found in Figure 4(d) (45/0.7/3).

In the Calendar Aging fitting process, the model has an average error of 0.8% between the estimated C_{loss} (dashed line) and the measured one (solid line), except for the condition $T = (55°C)$, $SoC = 5\%$, as the model does have an average error of 1.4% during this condition.

The results from the literature shows that during Calendar Aging, the LTO tends to degrade faster while the SoC is low compared to higher SoC, in addition to the effect of T. For Cycling Aging, the increase in T and

C-rate has significant effect on the LTO chemistry, and the degradation rate is faster when smaller SoC swings $\triangle SoC$ is applied.

Model Validation

In order to test the performance of the proposed model, it needs to be validated and compared to another known measured C_{loss} during cycling and Calendar Aging. For the Cycling Aging, the input data needed for the C_{loss} algorithm is extracted from (Baure & Dubarry, 2020), and the output is compared to the C_{loss} from the same reference. Table 1 and Table 2 first columns show a summary of the extracted data that is required for the Cycling Aging C_{loss}. Furthermore, both tables show the measured C_{loss} in 25/35°C, estimated C_{loss} and the error during 2500 equivalent cycles.

As shown in Table 1, the model does have an average error of 0.46% during Cycling Aging in 25°C, while it does have an average error of 1.39% in 35°C as shown in Table 2.

Table 1: Data validation during Cycling Aging measured in 25°C

Data cycling measured in 25 °C	Measured Capacity Loss %	Estimated Capacity Loss %	Error %
Median SoC = 15%, $\triangle$SoC = 5%	0.37	0.69	0.32
Median SoC = 50%, $\triangle$SoC = 5%	0.33	0.69	0.36
Median SoC = 85%, $\triangle$SoC = 5%	0.33	0.69	0.36
Median SoC = 15%, $\triangle$SoC = 45%	0.42	1.07	0.65
Median SoC = 50%, $\triangle$SoC = 45%	0.40	1.07	0.67
Median SoC = 85%, $\triangle$SoC = 45%	0.48	1.07	0.59
Median SoC = 50%, $\triangle$SoC = 75%	0.41	0.71	0.30

Table 2: Data validation during Cycling Aging measured in 35°C

Data cycling measured in 35 °C	Measured Capacity Loss %	Estimated Capacity Loss %	Error %
Median SoC = 15%, $\triangle$SoC = 5%	0.54	1.60	1.06
Median SoC = 50%, $\triangle$SoC = 5%	0.58	1.60	1.02
Median SoC = 85%, $\triangle$SoC = 5%	0.54	1.60	1.06
Median SoC = 15%, $\triangle$SoC = 45%	0.62	2.48	1.86
Median SoC = 50%, $\triangle$SoC = 45%	0.59	2.48	1.89
Median SoC = 85%, $\triangle$SoC = 45%	0.70	2.48	1.78
Median SoC = 50%, $\triangle$SoC = 75%	0.58	1.62	1.04

For the Calendar Aging, the required data is extracted and compared to the measured C_{loss} in (Dubarry et al., 2018). Table 3 shows a summary of the extracted data from this reference. To estimate the C_{loss} per 1 month,
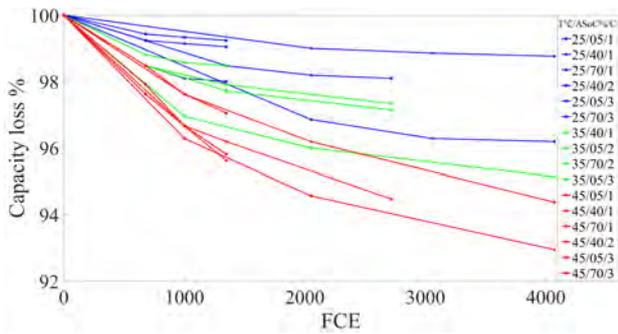
(a) Capacity Loss as a function of FCE
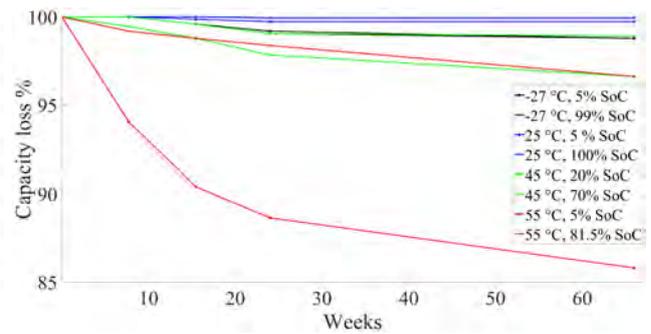
(b) Capacity Loss as a function of storage time

Figure 3: LTO measured Capacity Loss during Cycling and Calendar Aging (Dubarry & Devie, 2018).
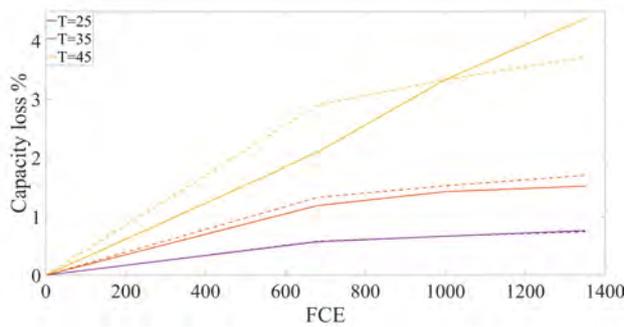

(a) Capacity Loss estimated output during a fixed $DoD(40\%)$
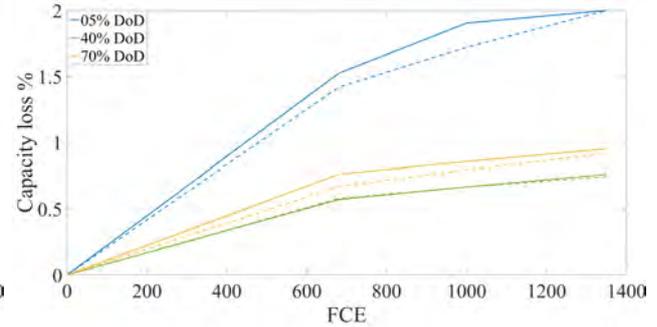
(b) Capacity Loss estimated output during a fixed $T(25C)$
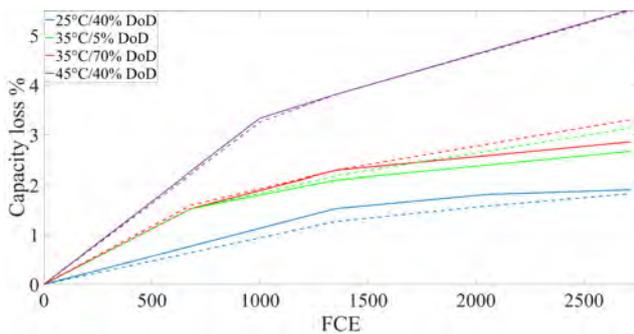
(c) Capacity Loss estimated output during a fixed $C-rate(2C)$
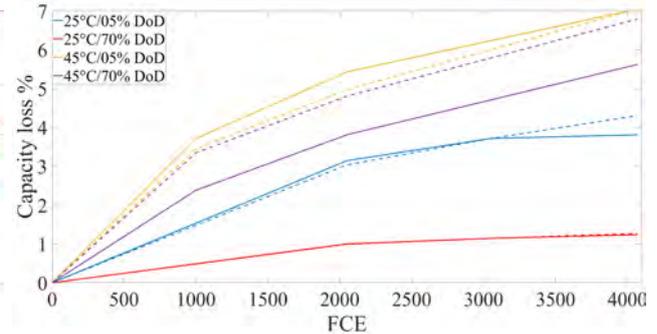
(d) Capacity Loss estimated output during a fixed $C-rate(3C)$

Figure 4: LTO estimated Capacity Loss during Cycling Aging.
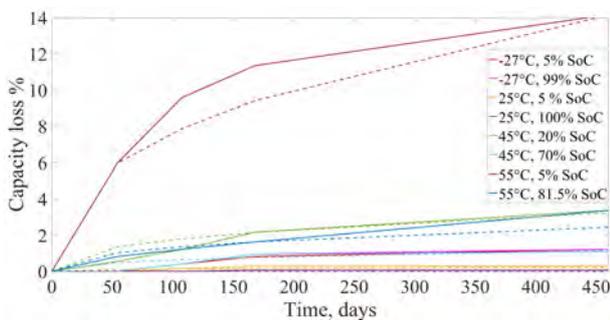
Figure 5: LTO estimated Capacity Loss during Calendar Aging.

the estimated C_{loss} and the measured one. The model does have an average error of 1.98% in Calendar Aging.

Table 3: Data validation during Calendar Aging

Temperature °C	SoC %	Capacity Loss (%/month)	Error %
-27	5	0.28	2.03
-27	99	0.28	1.95
25	50	0.05	0.67
25	100	0.05	0.68
45	20	0.76	2.74
45	70	0.24	1.76
55	5	3.97	3.37
55	81.5	0.73	2.67

Lithium iron phosphate battery (LFP)

Data Selection and Fitting

The measured data required for the fitting process during Cycling Aging is obtained from (Wang et al.,

`fitlm` from MATLAB is used to fit a linear regression model to obtain this value. Afterwards, the number of days is calculated and used as an input to the developed model. Table 3, 4th column shows the error between

(a) Capacity Loss as a function of FCE (Wang et al., 2011)

(b) Capacity Loss as a function of storage time (Naumann et al., 2018)
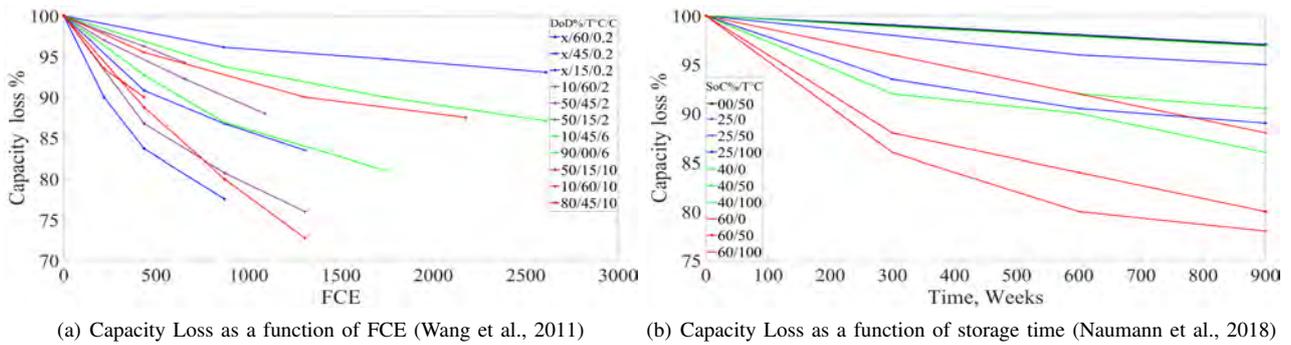
Figure 6: LFP measured Capacity Loss during Cycling and Calendar Aging.

(a) Capacity Loss estimated output during a fixed $DoD(50\%)$

(b) Capacity Loss estimated output during a fixed $T(60C)$

(c) Capacity Loss estimated output during a fixed $C-rate(0.5C)$

(d) Capacity Loss estimated output during a fixed $C-rate(10C)$
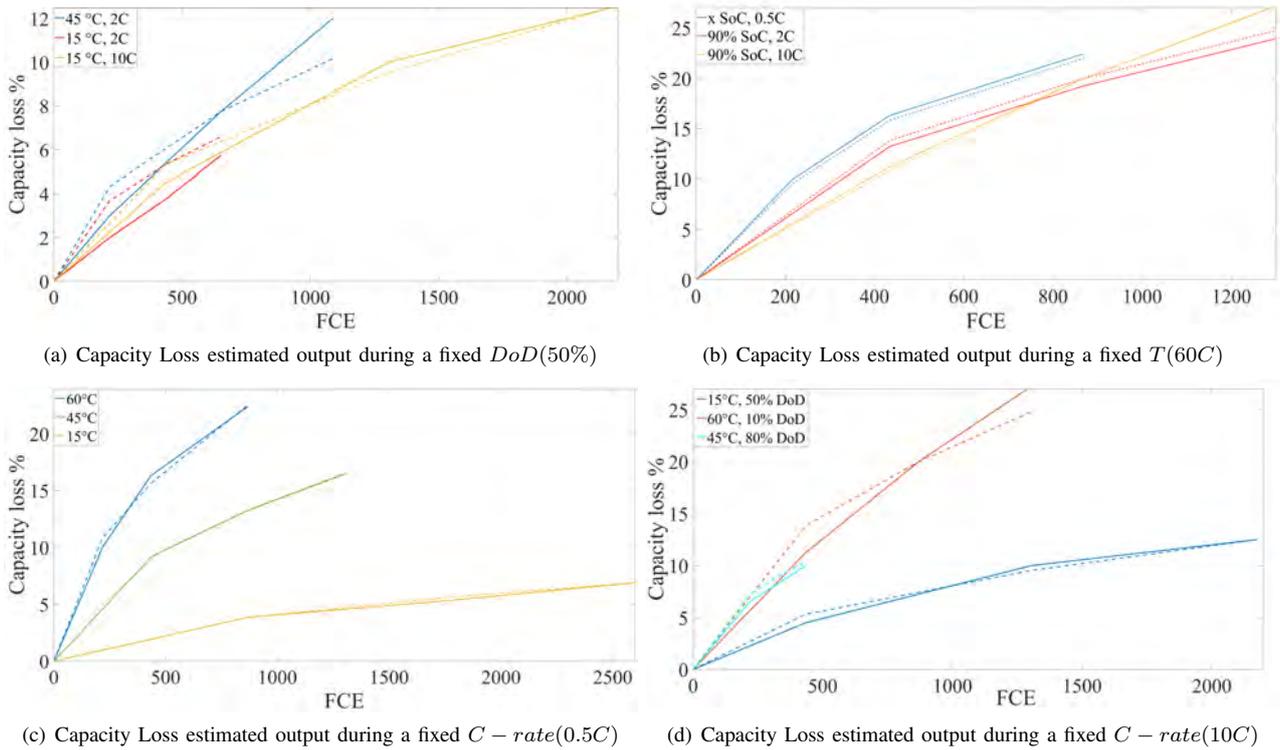
Figure 7: LFP Estimated Capacity Loss during Cycling Aging.
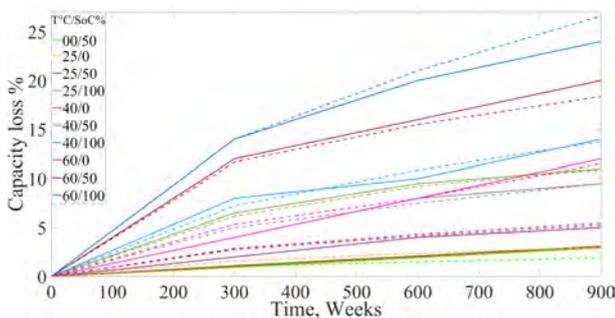
Figure 8: LFP Estimated Capacity Loss during Calendar Aging.

2011). In his work, he measured the C_{loss} during Cycling Aging in the following conditions; five different temperatures (0, 15, 25, 45, 60 °C), four levels of DOD (90%, 80%, 50%, and 10%), and four discharges rates (C/2, 2C, 6C and 10C). During discharge rate of C/2, the authors results showed that at such low rate, only temperature and FCE does have an effect on the C_{loss}, while DoD has a negligible effect on it. For the Calendar Aging C_{loss}, measured data was acquired from (Naumann et al., 2018), where the researchers studied the effect of different storage temperatures at the storage $SoC = 0\%, 50\%$, and 100%.

Afterwards, the fitting process for the Cycling Aging is carried on. Similar to the LTO chemistry, the fitting process is proceed in four different scenarios; During a fixed $DoD(50\%)$, fixed temperature (60°C), and Fixed C-rate (0.5C and 10C). Figure 7 shows the estimated C_{loss} (dashed line) compared to the measured one (solid line). The average error is found to be 0.43% at half the FCE for each point, and 0.54% at the total FCE.

During the Calendar Aging process, the model has an average error of 0.45% at 450 days between the estimated C_{loss} (dashed line) and the measured one (solid line), and 0.67% at 900 days.

From the literature, it can be concluded that during Calendar Aging, the LFP chemistry degradation rate remains constant with SoC at given T, which means the C_{loss} is affected more by T and for the Cycling Aging, the C_{loss} is strongly affect by T and t on the LFP chemistry, while the DoD has almost a negligible effect especially at low C-rate (0.5C). In contrast,

Model Validation

In this section, the model is validated by comparing the estimated C_{loss} output to other known measured C_{loss}. For Cycling Aging, data needed to estimate the C_{loss} and compare it to a measured one is obtained from (Marongiu et al., 2015). Table 4, first column shows a summary of the data that used to estimate the Cycling Aging C_{loss}. The value of the battery temperature is kept at 30°C, and FCE is ranged from 1700 up to 5000.

As can been seen in Table 4, the model estimated C_{loss} error ranges from 1.2% to 1.8%, with an average error of 1.15%.

To validate the model in Calendar Aging, the same method for obtaining the required data as been used with LTO is used, and data is extracted from (Dubarry et al., 2018) . Table 5 shows a summary of the extracted data from this reference and the calculated error between the estimated C_{loss} during Calendar Aging and the measured one. The model does have an average error of 1.83% in Calendar Aging.

Table 4: Data validation during Cycling Aging measured in 30°C

Data cycling measured in 30°C	Measured Capacity Loss %	Estimated Capacity Loss %	Error %
SoC=90%, 1C	19.7	18.5	1.2
SoC=50%, 1C	17.3	18.5	1.2
SoC=20%, 1C	6.2	7.6	1.4
SoC=90%, 3C	36.6	34.9	1.7
SoC=50%, 3C	27.9	26.6	1.3
SoC=20%, 3C	56.3	54.5	1.8
SoC=90%, 6C	15.1	13.8	1.3
SoC=50%, 6C	68.1	66.7	1.4
SoC=20%, 6C	38.6	37.1	1.5

Table 5: Data validation during Calendar Aging

Temperature°C	SoC %	Capacity Loss (%/month)	Error %
0	50	0.40	0.5
10	50	0.52	0.6
20	100	1.37	1.1
25	40	1.30	1.4
30	65	0.97	2.1
40	30	1.79	1.5
45	100	3.72	2.3
50	20	3.19	4.1
60	0	1.78	2.9

Application example

After presenting, identifying and validating the battery capacity degradation model using two different types of LIBs, an attempt to predict the capacity degradation using simulated duty cycle data for a *Mining*
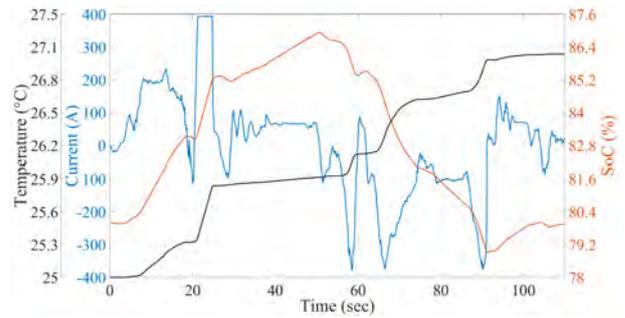

Figure 9: Simulated battery duty cycle data of a Mining Dumper.

Dumper based on (Immonen, 2003) work, is carried out. In this example, both LTO and LFP batteries are assumed to follow the same duty cycle, given in Figure 9 for Cycling Aging estimation, followed by 2 hours rest time during which the battery is cooled to the initial temperature (linear temperature decay) for the Calendar Aging estimation.

The battery system was simulated for 1000 repeated duty and rest cycles (described above) for both LTO and LFP chemistries. The required data for predicting the capacity loss, found in Figure 9, was then fed to the proposed SEM. The results show that, for the LFP chemistry, $C_{loss}^{cyc} = 3.1\%$, and $C_{loss}^{cal} = 8.31\%$. On the other hand, for the LTO chemistry, $C_{loss}^{cyc} = 5.2\%$ and $C_{loss}^{cal} = 2.22\%$, indicating a clear difference in the aging profiles of the two battery chemistries.

CONCLUSIONS

In this work, a SEM for estimating the C_{loss} of different LIBs during Cycling Calendar Aging for different operation condition has been developed. For modeling, different LIB measurement data is required for obtaining the models' fitting parameters. Afterwards, the model is validated by comparing the estimated C_{loss} to a known one. Results show that the proposed model is able to estimate the C_{loss} during Cycling and Calendar Aging of two different LIBs (namely LTO and LFP chemistries) with a high accuracy with a brief explanation of the effect of different parameters on the two LIBs. Furthermore, a simulated battery duty cycle of a Mining Dumper has been used to estimate the C_{loss} during Cycling and Calendar Aging of a Mining Dumper.

The proposed model can be used in case of the lack of the experimental data, where it can give out an acceptable accuracy while having low computation complexity and is simpler to implement in comparison to other model-based approaches. For instance, the developed SEM has an average error of 2% between the measured C_{loss} and the estimated C_{loss}, while the adaptive-models do usually have an average error less than 1%.

In the future, the proposed SEM framework should be validated by studying different battery chemistries. Another important topic left for future work is design optimization — optimal cell chemistry selection in par-

ticular — for HDBEVs, based on the proposed battery degradation models.

REFERENCES

Andre, D., Nuhic, A., Soczka-Guth, T., & Sauer, D. (2013, mar). "comparative study of a structured neural network and an extended kalman filter for state of health determination of lithium-ion batteries in hybrid electric vehicles". *Engineering Applications of Artificial Intelligence*, *26*(3), 951–961.

Baure, G., & Dubarry, M. (2020). "battery durability and reliability under electric utility grid operations: 20-year forecast under different grid applications". *Journal of Energy Storage*, *29*, 101391.

Berecibar, M., Garmendia, M., Gandiaga, I., Crego, J., & Villarreal, I. (2016). "state of health estimation algorithm of lifepo4 battery packs based on differential voltage curves for battery management system application". *Energy*, *103*, 784–796.

Bocca, A., Sassone, A., Shin, D., Macii, A., Macii, E., & Poncino, M. (2015). A temperature-aware battery cycle life model for different battery chemistries. In *Ifip/ieee international conference on very large scale integration-system on a chip* (pp. 109–130).

Cacciato, M., Nobile, G., Scarcella, G., & Scelba, G. (2016). Real-time model-based estimation of soc and soh for energy storage systems. *IEEE Transactions on Power Electronics*, *32*(1), 794–803.

Cen, Z., & Kubiak, P. (2020). Lithium-ion battery soc/soh adaptive estimation via simplified single particle model. *International Journal of Energy Research*, *44*(15), 12444–12459.

Dubarry, M., & Devie, A. (2018). "battery durability and reliability under electric utility grid operations: Representative usage aging and calendar aging". *Journal of Energy Storage*, *18*, 185–195.

Dubarry, M., Qin, N., & Brooker, P. (2018). "calendar aging of commercial li-ion cells of different chemistries–a review". *Current Opinion in Electrochemistry*, *9*, 106–113.

Grolleau, S., Delaille, A., Gualous, H., Gyan, P., Revel, R., Bernard, J., ... Network, S. (2014). Calendar aging of commercial graphite/lifepo4 cell–predicting capacity fade under time dependent storage conditions. *Journal of Power Sources*, *255*, 450–458.

Immonen, P. (2003). *Energy efficiency of a diesel-electric mobile working machine* (Unpublished doctoral dissertation). Lappeenranta University of Technology.

Kluschke, P., Gnann, T., Plötz, P., & Wietschel, M. (2019). "market diffusion of alternative fuels and powertrains in heavy-duty vehicles: A literature review". *Energy Reports*, *5*, 1010–1024.

Lagarias, J. C., Reeds, J. A., Wright, M. H., & Wright, P. E. (1998). Convergence properties of the nelder–mead simplex method in low dimensions. *SIAM Journal on optimization*, *9*(1), 112–147.

Liimatainen, H., van Vliet, O., & Aplyn, D. (2019). "the potential of electric trucks–an international commodity-level analysis". *Applied energy*, *236*, 804–814.

Liu, Z., Zhao, J., Wang, H., & Yang, C. (2020). "a new lithium-ion battery soh estimation method based on an indirect enhanced health indicator and support vector regression in phms". *Energies*, *13*(4), 830.

Marongiu, A., Roscher, M., & Sauer, D. U. (2015). "influence of the vehicle-to-grid strategy on the aging behavior of lithium battery electric vehicles". *Applied Energy*, *137*, 899–912.

Naha, A., Han, S., Agarwal, S., Guha, A., Khandelwal, A., Tagade, P., ... Oh, B. (2020, dec). "an incremental voltage difference based technique for online state of health estimation of li-ion batteries". *Scientific Reports*, *10*(1), 9526.

Naumann, M., Schimpe, M., Keil, P., Hesse, H. C., & Jossen, A. (2018). "analysis and modeling of calendar aging of a commercial lifepo4/graphite cell". *Journal of Energy Storage*, *17*, 153–169.

Petit, M., Prada, E., & Sauvant-Moynot, V. (2016, jun). Development of an empirical aging model for Li-ion batteries and application to assess the impact of Vehicle-to-Grid strategies on battery lifetime. *Applied Energy*, *172*, 398–407. doi: 10.1016/j.apenergy.2016.03.119

Schmalstieg, J., Käbitz, S., Ecker, M., & Sauer, D. U. (2014, jul). A holistic aging model for Li(NiMnCo)O2 based 18650 lithium-ion batteries. *Journal of Power Sources*, *257*, 325–334. doi: 10.1016/j.jpowsour.2014.02.012

Singh, P., Chen, C., Tan, C. M., & Huang, S.-C. (2019). Semi-empirical capacity fading model for soh estimation of li-ion batteries. *Applied Sciences*, *9*(15), 3012.

Standard practices for cycle counting in fatigue analysis (Standard). (2005). West Conshohocken, PA, USA: ASTM International.

Wang, J., Liu, P., Hicks-Garner, J., Sherman, E., Soukiazian, S., Verbrugge, M., ... Finamore, P. (2011). "cycle-life model for graphite-lifepo4 cells". *Journal of power sources*, *196*(8), 3942–3948.

Yang, K., Chen, Z., He, Z., Wang, Y., & Zhou, Z. (2020). "online estimation of state of health for the airborne li-ion battery using adaptive dekf-based fuzzy inference system". *Soft Computing*, *24*(24), 18661–18670.

AUTHOR BIOGRAPHIES

Mohammed Rabah PhD, is a Research Engineer at Computational Engineering and Analysis (COMEA) research group, Turku University of Applied Sciences, Finland.
E-mail:mohamed.rabah@turkuamk.fi

Eero Immonen D.Sc. (Tech.) and Adjunct Professor, leader of the Computational Engineering and Analysis (COMEA) research group at Turku University of Applied Sciences, Finland.
E-mail:eero.immonen@turkuamk.fi

Sajad Shahsavari works as Researcher at Turku University of Applied Sciences, and is a PhD student at University of Turku, Finland.
E-mail:sajad.shahsavari@turkuamk.fi

Mohammad-Hashem Haghbayan is a Post-Doctoral Researcher at Department of Future Technologies, University of Turku, Finland.
E-mail:mohammadhashem.haghbayan@utu.fi

Kirill Murashko PhD (M), is a Post-Doctoral Researcher at Department of Environmental and Biologic Science, University of Eastern Finland, Finland.
E-mail:kirill.murashko@uef.fi

Paula Immonen D.Sc. (Tech.) and Associate Professor of the School of Energy Systems, LUT University, Finland.
E-mail:paula.immonen@lut.fi

LogisticsLab: An academic software for decision-making in logistics

Mike Steglich

University of Applied Sciences Wildau
Hochschulring 1
15745 Wildau, Germany
mike.steglich@th-wildau.de

KEYWORDS

logistical decision-making, transportation problems, network flow problems, traveling salesman problems, Chinese postman problems, vehicle routing planning problems, location problems, problem-based learning and teaching

ABSTRACT

Logistical decision problems are a part of many courses in the field of logistics, management and operations research. It makes sense to illustrate these optimisation problems using case studies, which can be reproduced by students using suitable software. Often, solver add-ins in spreadsheets programs or general optimisation software are used, which on the one hand requires a high level of knowledge in Operations Research and on the other hand does not always allow an intuitive approach. This article describes the academic software LogisticsLab with which the distributors tie in with the idea of interactive decision support systems to systematically combine the experiences and intuitions of human decision-makers with the possibilities of computer-assisted modelling and optimisation of a wide range of logistical decisions.

INTRODUCTION

Students studying logistics management, supply chain management, etc. are faced with decision problems in transporting and storing of goods. Typical problems are transportation problems, network flow problems, traveling salesman problems, Chinese postman problems, vehicle routing planning problems and location problems (Rushton, Croucher, & Baker, 2017, p. 3ff). The students need to understand the problem itself and to learn how to formulate the problem mathematically, to solve it and to interpret the solution. A problem-based learning approach seems to be a suitable learning and teaching method for such tasks. Using real problems or realistic case studies, students should improve their problem-solving skills, their ability to independently acquire theoretical knowledge and develop critical thinking (Guimarães & Lima, 2021). In this context, suitable software is needed to solve the discussed problems.

Simulation software is an approach often used in Operations Management and Logistics and Supply Chain Management (LSCM) courses. A prominent example is the beer distribution game originally invented by Forrester in the 1960s (Dizikes, 2015) for which several software packages exist (Beergame.org, 2021). It deals with the dynamics within a supply chain with one homogeneous good including the bullwhip effect (Edali & Yasarcan, 2016). Other examples of teaching and learning approaches using simulation software for LSCM are Ștefan, Hauge, Hasse, & Ștefan (2019), Kanet & Stößlein (2008), Angolia & Pagliari (2018) and the Fresh connection simulation (Inchainge, 2021). In most of these simulations, students have to make and enter various decisions within the supply chain, although there seems to be no direct software support for this. Therefore, simulation approaches for these logistical decisions often also require an optimisation software package.

A frequently used approach to solving logistical decisions are spreadsheet programs with add-ins for optimisation (Winkels, 2012). However, this approach is more suitable for small problems due to the limitation of the invoked solvers and its size restrictions. Furthermore, the students must have a high level of knowledge in operations research techniques to formulate, solve and interpret the problems where such programs do not allow an interactive and intuitive way to do it (Mason, 2013).

Another option are optimisation environments, especially mathematical programming languages such as the commercial AMPL (AMPL, 2021), GAMS (GAMS, 2021), etc. or open-source packages like GMPL (GMPL, 2021), CMPL (CMPL, 2021) or Pyomo (Pyomo, 2021). There are also packages that implement routines for selected logistical decisions with mathematical languages. For example, Matlog implements routines for transport problems, vehicle routing and network optimisation in Matlab (Kay, 2016). Furthermore, the Google OR-Tools are very interesting, which offer application programming interfaces (APIs) for various programming languages for various optimisation problems including selected logistical decision problems (Google, 2021b). As with spreadsheet solver add-ins, students need a high level of knowledge in operations research, which is not always the case, especially in management courses. (Grasas & Ramalhinho, 2016, p. 378).

Another possibility is the use of commercial logistics software in university courses. It is obvious that the use of such software gives students a realistic insight into the management of supply chains. In addition, the acquired skills have a positive impact on their future position in the labour market (Campbell, Goentzel, & Savelsbergh, 2000). However, commercial logistics software is often too expensive to purchase and maintain (Sweeney, Campbell, & Mundy, 2010, p. 298f). Some manufacturers offer cheaper or free versions of their software for academic purposes, although these are often limited in their functionalities or the size of the problems. For example,

anyLogistix offers a commercial „supply chain analytics software to design, optimize and analyze a companies' supply chain" (anyLogistix, 2021a), for which a free version is available for academic purposes that does not include some of the commercial functionalities and is also limited in the number of customers and distribution centres and factories (anyLogistix, 2021b). If these limitations are not too restrictive for a particular course, then the use of free versions of commercial software makes sense. Another problem with commercial software packages can be that they do not offer the full range of decision problems discussed in university courses. A good example is Graphhopper (Graphhopper, 2021), which offers an excellent API for vehicle routing. However, it only supports vehicle routing. If other problems are to be discussed, additional software is needed.

Academic logistical software, which offers a wide range on supported logistical decision problems, is rather rare. As an example, the Toolbox of the Technical University of Dresden can be mentioned (Toolbox, 2013). In addition to problems in procurement logistics and production logistics, shortest-path problems, network flow problems, transport problems, vehicle routing problems and location problems can also be solved. However, the software is only available in German, seems rather orientated on techniques and does not support all relevant aspects of logistical decision making. Another example is a web-based decision support system for vehicle routing problems developed by Grasas & Ramalhinho (Grasas & Ramalhinho, 2016). Embedded in a problem-based learning approach, the tool is intended "to better understand the solution method and to show students the need for decision-making software in complex problems" (Grasas & Ramalhinho, 2016, p. 389). The combination of the easy-to-use web-based tool and the didactic approach is very interesting. However, this tool can only be used for vehicle routing problems.

The motivation for the development of LogisticsLab was, considering all the possibilities of using software for teaching logistical decision problems, to create an academic software that is easy and intuitive to use and offers a wide range of supported logistical decision problems.

This paper describes LogisticsLab. In the first section, an overview of the software including the main functionalities and the file formats are given. The following section is intended to illustrate how students can solve realistic case studies by using LogisticsLab. The described case studies have been used in several courses at Bachelor level in German and international universities. The paper ends with a summary.

LOGISTICSLAB

Overview

LogisticsLab is an academic software for decision support in logistics. This software can be classified as a problem-oriented decision support system that can be used to model and solve transportation problems, network flow problems, traveling salesman problems, Chinese postman problems, vehicle routing planning problems and location problems in different variations.

LogisticsLab was originally developed in the 1990s under the name *EUS-Lehrsoftware* by Dieter Feige for logistics courses at the Friedrich-Alexander University of Erlangen-Nuremberg and fundamentally revised and expanded by the author of this article.

LogisticsLab is provided under the LogisticsLab Academic License. It is available for academic use (such as research and teaching or for reproducing case studies in academic textbooks), free of charge and without warranty. It can be downloaded at http://logisticslab.org and is available for Microsoft Windows. No installation is required. The user only has to unzip the downloaded zip file and the contained executables can be used immediately.

Main functionalities

After unpacking the downloaded distribution file, the LogisticsLab folder contains several binary files with specific names that provide the following functionalities.

TPP

LogisticsLab/TPP is a software that can be used to solve classical transport problem (Hillier & Lieberman, 2015, p. 319ff) and bottleneck transport problems (Garfinkel & Rao, 1971). Minimisation and maximisation (or Mini-Max and MaxiMin for bottleneck problems) can be selected for the objective function sense. It is also possible to invoke step-fixed costs for source-destination-relations. In addition, various constraints can be taken into account (capacities, single-sourcing, supply or demand surplus, etc.).

NWF

LogisticsLab/NWF is intended for solving network flow problems, which can be modelled as min-cost flow problems (Ghiani, Laporte, & Musmanno, 2013). One of the most important problems to be solved with this model is the transhipment problem (Hillier & Lieberman, 2015, p. 401).

TSP

LogisticsLab/TSP is used for solving traveling salesman problems (Applegate, Bixby, Chvátal, & Cook, 2006) and open traveling salesman problems (Steglich, Feige, & Klaus, 2016, p. 312ff) for symmetrical and asymmetrical distances between the nodes to be approached. For open traveling salesman problems, four different types can be specified (fixed start and end node, fixed start node and free end node, fixed end node and free start node, free start and end node). Symmetric distances can be calculated as airline distances using the coordinates of the nodes as Euclidian distances, Manhattan distances and great circle distances. It is also possible to define directed or undirected edges with given distances. In addition, a distance matrix containing distances obtained by APIs of geographic information systems like Google Maps (Google, 2021a), Bing Maps (Microsoft, 2021) or

OpenStreetMap (OpenStreetMap, 2021b) can be imported via LogisticsLab's file format.

CPP
LogisticsLab/CPP is intended for solving Chinese postman problems in undirected, directed and mixed networks. Starting from a start node, each edge is to be traversed at least once and returned to the start node at the end of the round trip. The edge-oriented round trip to be found should minimise the total distance (or other adequate evaluation) (Steglich, Feige, & Klaus, 2016, p. 320ff).

VRP
LogisticsLab/VRP is a software with which vehicle routing problems in different variants can be solved. The capacitated vehicle routing problem (VRP) aims to find a set of routes at a minimal cost level, beginning and ending the routes at a depot, so that the demands of all nodes (customers, cities, etc.) are satisfied. Each node is visited at least once and each vehicle has a limited capacity (Williams, 2013, S. 198ff). With VRP, a user is able to solve such problems in symmetric graphs using airline distances (Euclidian, Manhattan and great circle distances), distances defined for edges or distances obtained by geographic information systems (e.g. Google Maps, Bing Maps, OpenStreetMap). Several additional constraints can be invoked into the decision (route duration, route lengths, etc.)

WLP
Warehouse Location Problems are discrete location problems where both fixed location costs and variable transport costs as well as the capacities of the potential sources and the demands of the destinations are included in the location decision. The objective is to minimise the sum of the transportation costs and the fixed costs of building and running warehouses by deciding which warehouses are established and which customer is delivered to by which warehouse (Fernández & Landete, 2015, S. 47ff). LogisticsLab/WLP is designed as a software for solving single-stage capacitated and uncapacitated Warehouse Location Problems, always assuming that a demand node can only be supplied from one location at a time (single-sourcing).

CLP
LogisticsLab/CLP is a software to solve continuous (and discrete) median and centre problems. Continuous location problems determine the best possible greenfield location. The point where a planned logistics node can be optimally placed is to be found within a certain area on a flat or spherical surface. In contrast, in discrete location problems, only existing logistics nodes can act as locations (Steglich, Feige, & Klaus, 2016, p. 394ff). The following two approaches are distinguished: The *median problem* aims at minimising the sum of the (weighted) distances between the sources and the destinations and consists in finding the optimal locations of the sources and an optimal assignment of the destinations to the sources. Depending on the number p of locations sought, these problems are also called p-median problems (Neema, Maniruzzaman, & Ohgai, 2011). The *centre problem* aims at minimising the maximum (weighted) distance over all relations between the locations of the sources and the assigned destinations. Depending on the number p of locations to be determined, these problems are also called p-centre problems (Calik, Labbé, & Yaman, 2015).

User Interface

The user interface of each part of LogisticsLab follows a uniform structure, as shown in Figure 1 for LogisticsLab/VRP as an example. In addition to the usual elements, the user interface has a network area on the left-hand side of the application window with the graphical representation of the network and a data area on the right-hand side in which several areas can be selected via tabs and filled with the necessary data. Generally, interactive data entry and modification takes place in the data area, but it is also possible to modify some of the data in the network area interactively. The network and the data area can be spread over the entire window.

File Formats

The data of the programs are stored in tab separated text files with a specific format. To distinguish the data for the different programs, the files are given the program abbreviation as file extensions, i.e.: TPP, NWF, TSP, CPP, VRP, WLP and CLP.

The following principles apply to data files:
- Each file begins with the keyword for the program for identification. This keyword can be followed in the same line by information on the date and time.
- The next line contains a comment.
- For some of the applications, a section for options follows. This begins with the keyword OPTIONS, after which the corresponding options are listed in the following lines.
- The problem data is divided into data sections. Each section begins with a line in which the corresponding keyword (e.g., SOURCES, DESTINATIONS) for the respective data must appear. In most cases, the size specifications for the data section (e.g., number of sources and destinations) follow in a subsequent line. The following lines contain the corresponding data. The records are defined line-wise as ASCII text. The data entries are separated with tabulators.
- The end of file is expressed with the keyword EEE in the last line.

The data files can be created via the menu function *Save Problem* or *Save Problem as*. However, it is also possible to edit and maintain an existing data file (as well solution files) with a spreadsheet software (e.g., Microsoft Excel or LibreOffice/Calc). Finally, the data must be saved as a pure text file with data separation by tabs.
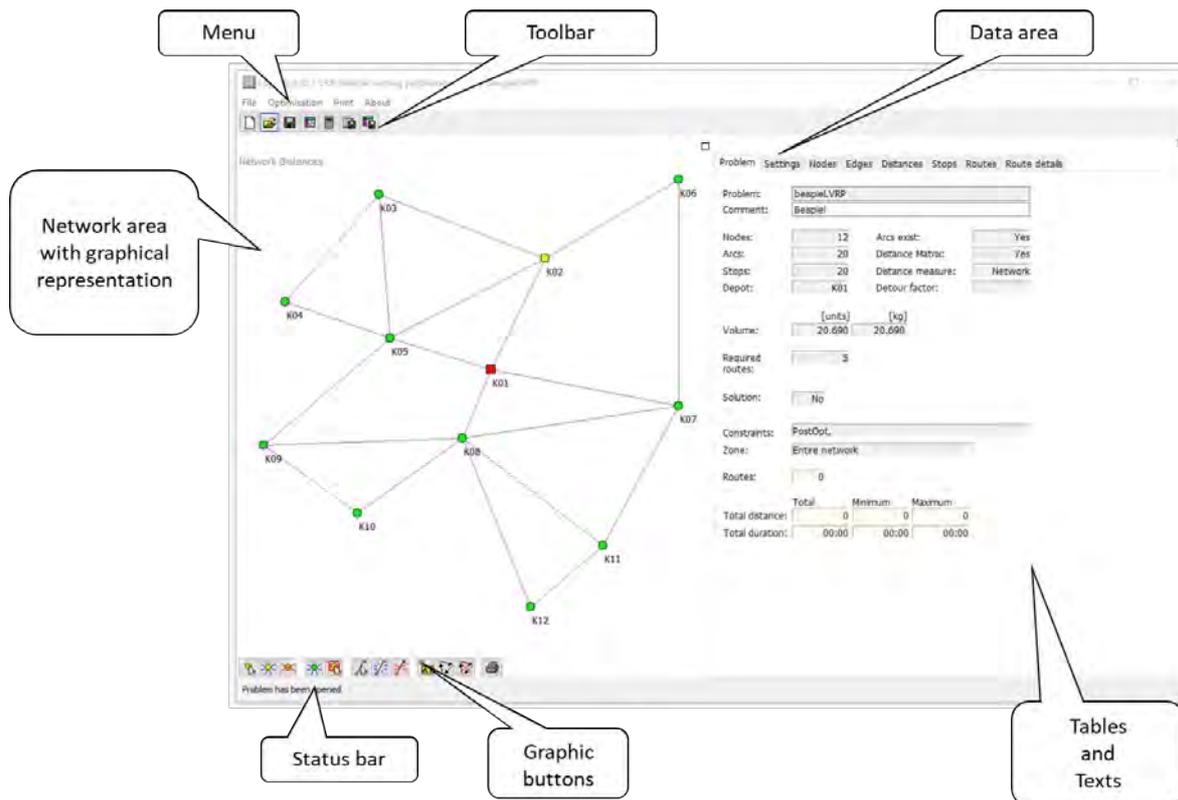
Figure 1: User interface of LogisticsLab/VRP

SELECTED CASE STUDIES

This section is intended to illustrate how students can solve realistic case studies by using LogisticsLab. Both case studies have been used in courses at Bachelor level at several European universities, often using a problem-based learning and teaching approach (Hung, 2011) (Grasas & Ramalhinho, 2016).

In a first step, the students, organised in groups, have to analyse a given realistic problem and work out the theoretical basis of the problem in order to formulate a suitable mathematical model. The mathematical model helps the students to understand and formulate the objective and the constraints of the problem and what data is needed to solve the case. In addition, students have to obtain the problem data from various sources. Afterwards, the problem, including the obtained data, has to be entered into LogisticsLab and solved with it. The last but not least task is the interpretation and presentation of the solution found.

Street condition survey in Manhattan

In this case study, the students were asked to find the shortest route to examine the condition of part of the streets in Manhattan, NY. As shown in Figure 2, it is the area bounded on the northwest by the 10th Avenue/West 57th Street intersection and on the southeast by the 6th Avenue/East 42nd Street intersection. All streets in this area are to be driven at least once.

This problem can be solved as a Chinese postman problem which can be understood as an edge-oriented round-trip problem. Consider a network $G = (N, E)$, where N is a set of nodes and E is a set of edges joining pairs of nodes. Each edge has to be used at least once, whereby the tour starts and ends at a specific node. The total distance of this route is to be minimised (Ahuja, Magnanti, & Orlin, 2013, S. 740) (Steglich, Feige, & Klaus, 2016).
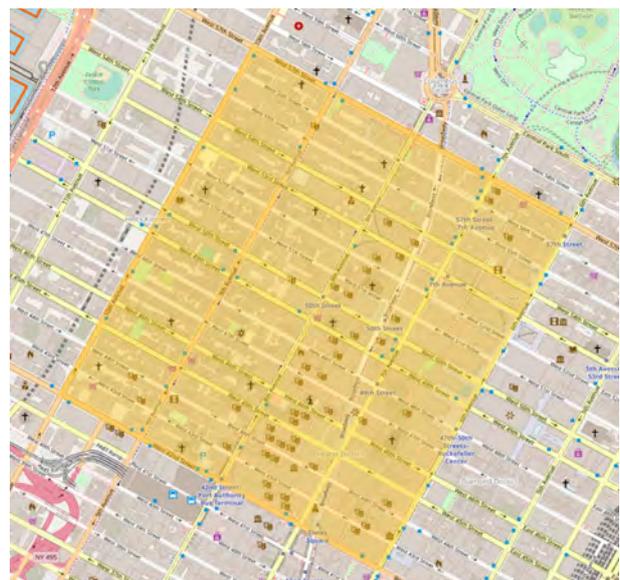
Figure 2: Area in Manhattan (OpenstreetMap, 2021a)

In this case, the nodes are defined by the street intersections and the edges by the streets, considering the directions of the streets. The coordinates of the nodes and the distances of the edges can be obtained by using the APIs of geographical information systems like Google Maps, Bing Maps or OpenStreetMap.

After starting the application, the user interface shown in Figure 3 appears, which, like all other components of LogisticsLab, consists of a network area and a data area.
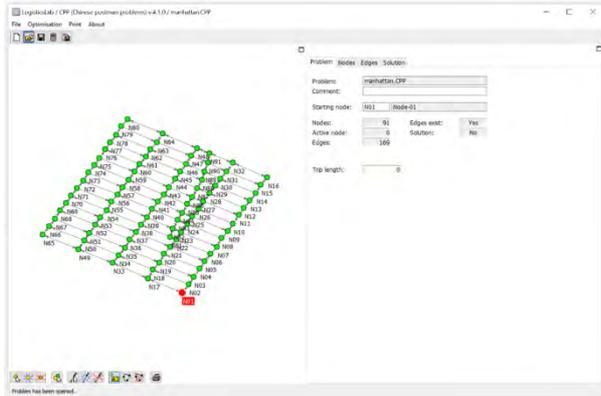
Figure 3: User interface of LogisticsLab/CPP

The left side of the CPP window is used to visualise the network graph depending on the data entered. If the network graph contains connections between the nodes, these are shown as edges. Directed edges are represented by arrows. Undirected edges do not have arrowheads. The data area on the right-hand side of the CPP user interface contains four worksheets for entering data and outputting optimisation results (*Problem*, *Nodes*, *Edges*, *Solution*).

The first step using LogisticsLab/CPP is to generate a new problem. To do this, the menu item *File → New Problem* or the *New Problem* button in the toolbar is selected. In addition to a comment, the number of *nodes* (in this case 91 nodes) and the maximum distance between the nodes (*Max. distance)* must be entered (Figure 4). The latter is used when generating randomly based coordinates (*Coordinates → Randomly*). The generated coordinates can be edited subsequently.
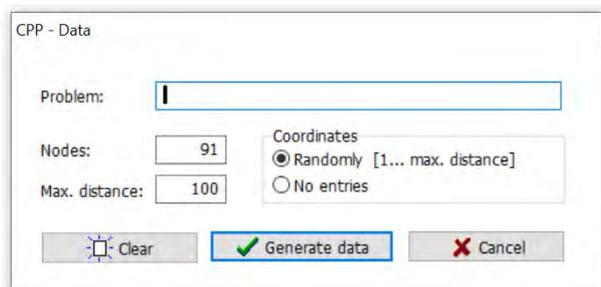
Figure 4: Creating a problem in LogisticsLab/CPP

After generating the problem, the data can be entered in the data area. The nodes can be edited in the *Nodes* tab (Figure 5).

A user can either enter the number of nodes or add each node individually to the list with the *Add* button.
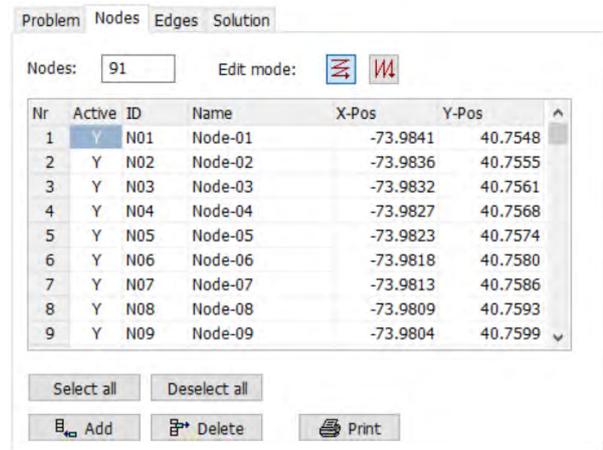
Figure 5: *Nodes* tab in LogisticsLab/CPP

Each node is specified by a selection (*Active* - selection for a calculation), a unique node ID (*ID*), a node name (*Name*) and an X and a Y coordinate. Geographical coordinates can also be used instead of simple X and Y coordinates. In this case, the longitude must be entered as the X-value and the latitude as the Y-value.

An easier way is to open the generated CPP file in a spreadsheet software, edit the names and copy and paste the coordinates found with Google Maps, Bing Maps or OpenStreetMap into the file (Figure 6).

Figure 6: Editing a CPP file in Microsoft Excel

The edges have to be defined in the *Edges* tab (Figure 7). To do this, the node from which the edges start must be specified in the *From node* field. The edges are defined in the *To nodes* list by entering the name, length and type of the edge. The type can be *U* for an undirected edge or *D* for a directed edge.
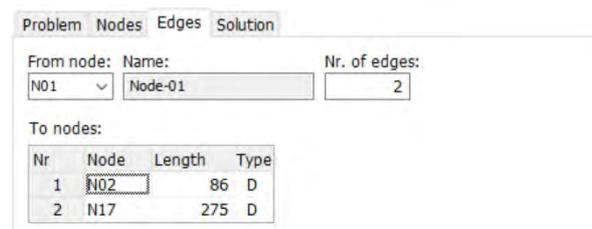

Figure 7: *Nodes* tab in LogisticsLab/CPP

For large problems, it is also recommended for the edges to open the CPP file in a spreadsheet program and copy and paste the distances into it.

After entering all the data, the problem can be solved by selecting either the *Optimisation → Start Optimisation* menu or the *Optimise* button in the toolbar. To solve a Chinese postman problem, the following two steps are carried out by LogisticsLab/CPP automatically (Steglich, Feige, & Klaus, 2016, p. 320 ff):

1. Cost- or distance-minimal extension of the network into a Euler network,
2. Determination of the Euler tour.

The cost- or distance-minimal extension of the existing network is solved on the basis of an integer linear optimisation model (Winkels, 2012, p. 590f) (Steglich, Feige, & Klaus, 2016, p. 325f), the results of which are used by a recursive algorithm to determine an Euler tour (Mattfeld & Vahrenkamp, 2014, p. 223f) (Steglich, Feige, & Klaus, 2016, p. 322ff) and thus to solve the Chinese postman problem.

After solving the problem, the graphical representation of the solution appears in the *Network* area (Figure 9) and the details in the *Solution* tab (Figure 8). In the Network area (Figure 9), the thickness of the edges indicates whether they have to be traversed several times.

Figure 8: *Solution* tab in LogisticsLab/CPP
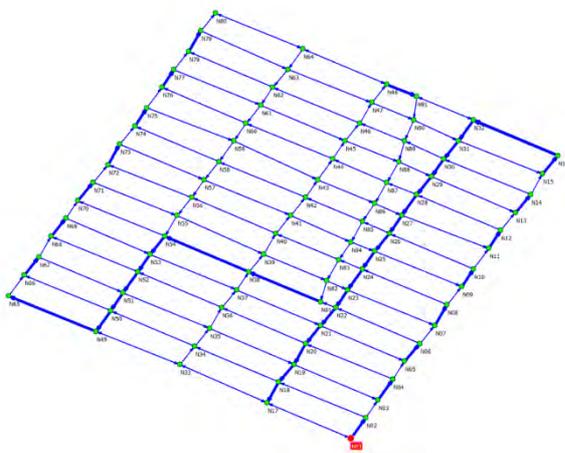
Figure 9: *Network area* in LogisticsLab/CPP

The best solution found has a length of 31.886 km (Figure 8). Both the graphical solution and the list of the entire route can be printed. It is also possible to save the solution in a solution file which can easily be imported into a spreadsheet program.

Planning three disaster relief centres in the state of Brandenburg, Germany

In this case study, students were asked to find the optimal locations of three disaster relief centres for the state of Brandenburg. The task of the centres is to send technology and emergency forces to the towns and municipalities as fast as possible in the event of a disaster. The maximum distance of the centres to the potential places of operation is to be minimised. Since cities and municipalities with many inhabitants are to be given priority in the event of a disaster, the distances are to be weighted according to the number of inhabitants. The coordinates and the number of inhabitants of the cities and municipalities can be taken from data of the German Federal Statistical Office (Statistisches Bundesamt, 2020). The locations of the disaster relief centres have to be planned on greenfield sites. The problem is a continuous *p*-centre problem without considering capacities and with demand-weighted distances, which can be formulated as follows (Drezner, 1984) (Drezner, 2011):

$$R \rightarrow min! \tag{1}$$
$$s.t.$$
$$b_j \cdot x_{ij} \cdot d\big((u_i, v_i), (x_i, y_i)\big) \leq R \quad ; i \in S, j \in D \tag{2}$$
$$\sum_{i \in S} x_{ij} = 1; j \in D \tag{3}$$
$$x_{ij} \in \{0,1\} \quad ; i \in S, j \in D \tag{4}$$

Indices and sets
- $i \in S$ sources
- $j \in D$ destinations

Parameters
- b_j demand of destination j
- x_j, y_j coordinates of destination j

Function
- $d(\cdot)$ distance function

Variables
- R maximum demand weighted distance
- x_{ij} assignment variable whether source i serves destination j
- u_j, v_j coordinates of source i

According to the expressions (1) and (2), the maximum demand-weighted distance R between the sources and the destinations is to be minimised. These distances are to be determined for all combinations of sources and destinations with a suitable distance function $d(\cdot)$. These distances are only relevant if the corresponding assignment variables x_{ij} are equal to the value one, i.e., destination j is assigned to source i. Expression (3) ensures that a destination is assigned to exactly one source (Steglich, Feige, & Klaus, 2016, p. 411).

The problem can be solved with LogisticsLab/CLP, which, like all other parts of LogisticsLab, has a user interface consisting of a network area and a data area (Figure 10).
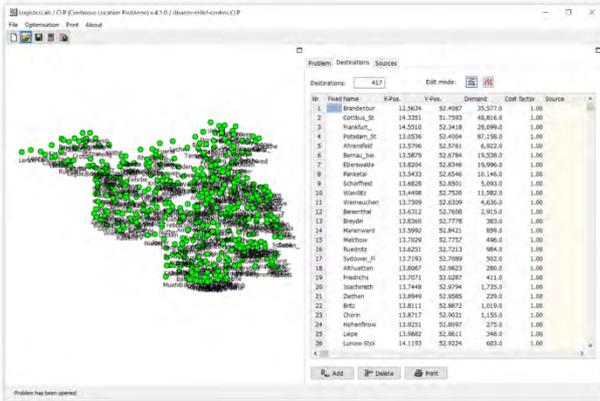

Figure 10: User interface of LogisticsLab/CLP

The first step is to create the problem file and to collect the problem data. Since the state of Brandenburg consists of 417 towns and municipalities, a problem with 417 destinations is to be created via the menu item *File → New Problem* or the *New Problem* button in the toolbar, as shown in Figure 11.
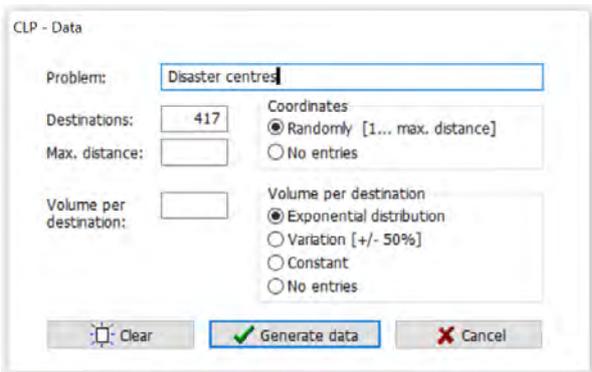
Figure 11: Creating a problem in LogisticsLab/CLP

Afterwards, the data of the generated problem can be entered in the data area (Figure 12). For all destinations, the names, the coordinates and the demands have to be entered, whereby for this problem the demands are defined by the number of inhabitants of the corresponding town or municipality. Since geographical coordinates are used, the longitude must be entered as the X-value and the latitude as the Y-value. The names, the coordinates and the inhabitants of the towns and municipalities can be taken from an Excel file of the German Federal Statistical Office (Statistisches Bundesamt, 2020). However, it is recommended to copy the mentioned data from this Excel file and paste it into the CLP file opened in a spreadsheet program (Figure 13).
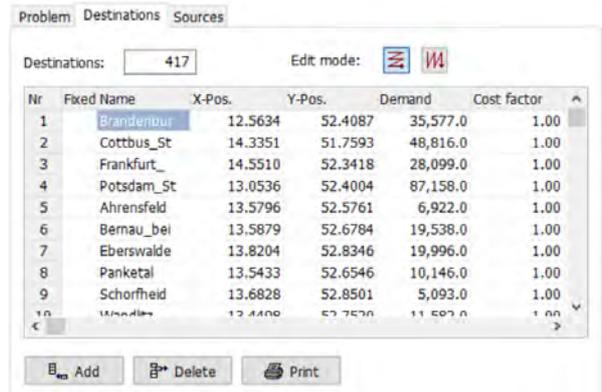

Figure 12: *Destination* tab in LogisticsLab/CLP
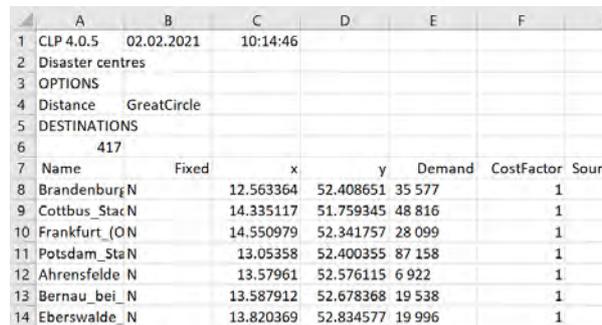
Figure 13: CLP problem file in Excel

The *Sources* tab (Figure 14) contains the details of the locations whose positions are to be determined by the optimisation. By entering the number of sources, the size of the input sheet is adjusted automatically. The names and the capacities (if needed) can be edited in this tab.
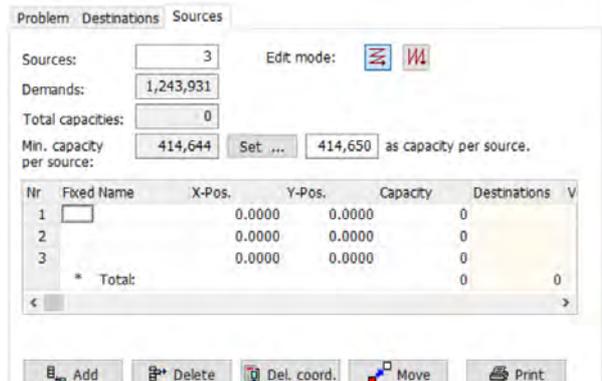

Figure 14: *Sources* tab in LogisticsLab/CLP

Once all the data has been entered, the problem can be solved. To do this, the problem type *centre of gravity* and the distance function to be used (*Distance → Great Circle*) must be selected in the *Problem* tab (Figure 15). Since the number of inhabitants of the destinations influence the priority in the event of a disaster, it must be specified that the distances are weighted with the demands of the destinations (checkbox *Demand weighted*).
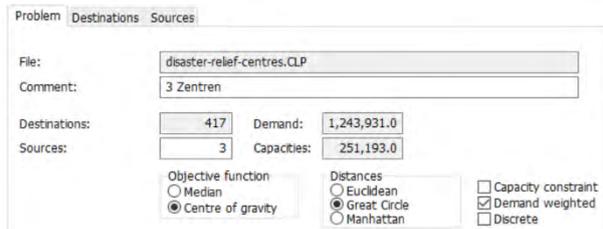
Figure 15: *Problem* tab in LogisticsLab/CLP

Optimisation is started by selecting either the *Optimisation → Start Optimisation* menu or the *Optimise* button in the toolbar. LogisticsLab uses state of the art heuristics to solve continuous location problems. As is usual for heuristics, the solutions may vary from run to run.

The solution of the problem is displayed in both the network area and data area. In the network area (Figure 16), the locations found for the three disaster relief centres and additionally the assignments of the towns and municipalities as potential places of a disaster event are shown. The exact locations of the disaster relief centres can be found in the *Sources* tab (Figure 17) in the columns *X-Pos* and *Y-Pos*. Besides to the number of assigned destinations (*Destinations*), the nearest destinations to the greenfield locations of the sources are shown.
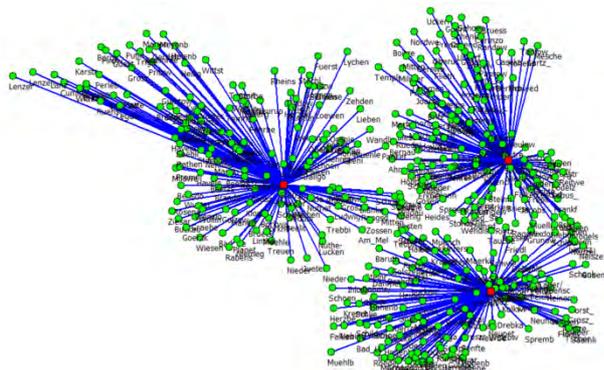

Figure 16: Network area in LogisticsLab/CLP

Figure 17: Solution in the *Sources* tab

The objective was to minimise the maximum of the demand-weighted distances between the locations to be found and the destinations. As shown in Figure 18, the objective function value equals 922,193.6 person-kilometres (similar to passenger-distance). This value is the minimum *Maximum Costs* of all relations between the new locations of the disaster relief centres and the assigned destinations. It results from the multiplication of the number of the inhabitants (*Demand*) and the distances. The cost factor is set to the value one and can be

ignored. This indicator is reasonable because it is assumed that an increasing number of inhabitants also leads to the increase of the amount of technology and emergency forces to be sent to the location of a disaster event.

Figure 18: Key indicators in the *Problem* tab

The assignments of the destinations to the newly found locations of the sources can be found in the *Destinations* tab (Figure 19). The relation with the maximum inhabitant weighted distances is between the destination *Bernau* (Nr. 6) and the first source which is located nearby *Neuhardenberg* (Figure 19 and Figure 17).

Figure 19: Solution in the *Destination* tab

The objective of this case study can be changed if it is assumed that the amount of technology and emergency forces is the same for all events and therefore independent of the number of inhabitants. In this case, only the maximum over all distances between the locations and the potential disaster event places is to be minimised. This can be achieved simply by not activating the option *Demand weighted* in the *Problem* tab (Figure 15).

After running the optimisation again, a new solution can be found where the locations of the three disaster relief centres are not attracted to destinations with high populations. This can be seen in the graphical solution in the network area (Figure 20) and on the newly determined coordinates of the sources in the *Sources* tab (Figure 21).

Since a heuristic is used, the solutions may be different for different optimisation runs. The minimal maximum distance found equals 73.2 kilometres (Figure 22). That means that the teams have to travel a maximum of 73.2 kilometres over all relations between the newly determined locations of the disaster relief centres and the potential places of a disaster event. This is less than the minimal maximum distance of the original problem which is equal to 128.5 kilometres (Figure *18 → Distances → Maximum*).
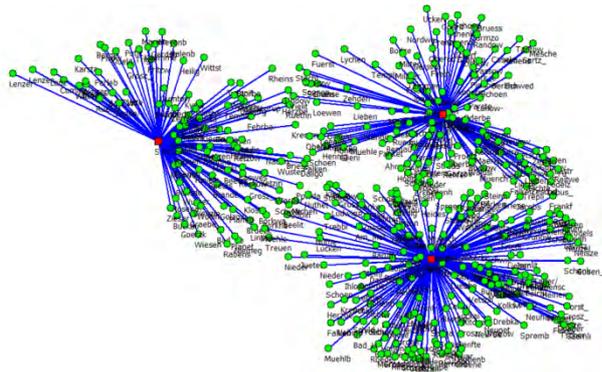
Figure 20: Graphical solution of the unweighted case in the network area in LogisticsLab/CLP

Figure 21: New locations of the unweighted case in the *Sources* tab in LogisticsLab/CLP

Figure 22: Key indicators in the *Problem* tab for the unweighted case

However, the locations found with a continuous p-centre problem cannot be implemented in reality in every case. So, it could be possible that the three locations nearby *Havelaue*, *Schoenwalde* and *Hohenfinow* (Figure 21) cannot be realised for many reasons. If only existing nodes in a network are available as locations for the sources, a discrete *p*-centre problem has to be solved.

In difference to a continuous problem, the coordinates of the locations are not variables, because only predefined nodes of the set S of potential sources can be used as locations. Additionally, all distances between the potential sources $i \in S$ and the destinations $j \in D$ are known. Regarding the mathematical model, please refer to the literature (Calik, Labbé, & Yaman, 2015, p. 83f).

LogisticsLab/CLP is mainly intended to solve continuous *p*-median and *p*-centre problems. However, it offers also to solve both as discrete location problems using a heuristic. To do so, a user only has to choose the option *Discrete* in the *Problem* tab (Figure 15) and to start the optimisation again. After selecting this option, a new solution can be found for the given unweighted case. It

should also be mentioned for the discrete problem that the solutions of different optimisation runs may differ due to the non-deterministic nature of the underlying heuristic. The discrete solution is shown as graphical solution (Figure 23) in the network area and the coordinates are shown in the *Sources* tab (Figure 24).
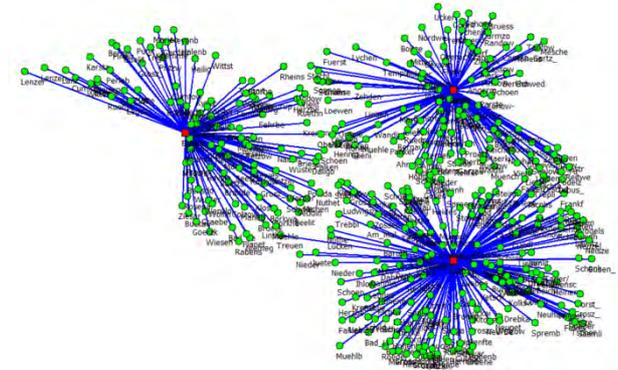
Figure 23: Graphical solution of the discrete case in the network area in LogisticsLab/CLP

It is interesting to note that the new locations for the disaster relief centres in *Ziethen*, *Luebben* and *Havelaue* are partially unequal to *Havelaue*, *Schoenwalde* and *Hohenfinow*, which are the closest destinations to the locations found for the continuous problem. Moving the continuous locations to the closest destinations may lead to suboptimal solutions.

Figure 24: New locations of the discrete case in the *Sources* tab in LogisticsLab/CLP

The minimal *Maximum Distance* found for the discrete problems is equal to 75.6 kilometres (Figure 25). This is only 2.4 kilometres more for this more realistic solution compared to the solution of the continuous problem.

Figure 25: Key indicators in the *Problem* tab for the unweighted case

SUMMARY

Students who are confronted with logistical decision problems in several courses need appropriate software to solve and interpret realistic case studies. The frequently used solver add-ins in spreadsheet programs or general optimisation software require a high level of knowledge in operations research and do not always allow intuitive access. Commercial logistics software and the rather rare academic software often offer only a smaller selection of supported logistical decisions.

This article described the academic software LogisticsLab, which supports a wide range of logistical decisions. The idea of the software is the systematic combination the experiences and intuitions of human decision-makers with the opportunities of computer-assisted modelling and optimisation in the form of an interactive decision support system. Two realistic case studies, used in courses at several European universities, demonstrate how LogisticsLab can be used to gain a better understanding of logistical decision-making in an interactive and intuitive way.

REFERENCES

Ștefan, I. A., Hauge, J. B., Hasse, F., & Ștefan, A. (2019). Using Serious Games and Simulations for Teaching Co-Operative Decision-making. *7th International Conference on Information Technology and Quantitative Management (ITQM 2019)* (pp. 745-753). Procedia Computer Science.

Ahuja, R. K., Magnanti, T. L., & Orlin, J. B. (2013). *Network Flows: Theory, Algorithms, and Applications.* Harlow: Pearson New International Edition.

AMPL (2021). Retrieved February 2021, https://ampl.com.

Angolia, M. G., & Pagliari, L. R. (2018). Experiential Learning for Logistics and Supply Chain Management Using an SAP ERP Software Simulation. *Decision Sciences Journal of Innovative Education, 16*(2), 104-125.

anyLogistix (2021a). *anyLogistix - Features overview.* Retrieved April 2021, https://www.anylogistix.com/alx-features/.

anyLogistix (2021b). *anyLogistix - Personal Learning Edition.* Retrieved April 2021, https://www.anylogistix.com/personal-learning-edition/.

Applegate, D., Bixby, R., Chvátal, V., & Cook, W. (2006). *The traveling salesman problem: A Computational Study.* Princeton: Princeton University Press.

Beergame.org (2021). *Versions.* Retrieved April 2021, https://beergame.org/the-game/versions/.

Calik, H., Labbé, M., & Yaman, H. (2015). p-Center Problems. In G. Laporte, S. Nickel, & F. Saldanha da Gama (Eds.), *Location Science* (pp. 79-92). Cham et al.: Springer.

Campbell, A., Goentzel, J., & Savelsbergh, M. (2000). Experiences with the use of supply chain management software in education. *Production and Operations Management, 9*(1), 66-80.

CMPL (2021). Retrieved February 2021, http://coliop.org.

Dizikes, P. (2015). *The Many Careers of Jay Forrester.* Retrieved April 2021, https://www.technologyreview.com/2015/06/23/167538/the-many-careers-of-jay-forrester/.

Drezner, Z. (1984). The p-Centre Problem-Heuristic and Optimal Algorithms. *The Journal of the Operational Research Society, 35*, 741–748.

Drezner, Z. (2011). Continuous Center Problems. In H. Eiselt, & V. Marianov (Eds.), *Foundations of Location Analysis* (pp. 63–78). Springer US.

Edali, M., & Yasarcan, H. (2016). Results of a beer game experiment: Should a manager always behave according to the book? *21*(S1), 190-199.

Fernández, E., & Landete, M. (2015). Fixed-Charge Facility Location Problems. In G. Laporte, S. Nickel, & F. Saldanha da Gama (Eds.), *Location Science* (pp. 47-78). Cham et al.: Springer.

GAMS (2021). Retrieved February 2021, https://www.gams.com.

Garfinkel, R., & Rao, M. (1971). The bottleneck transportation problem. *Naval Research Logistics Quarterly* (18), 465–472.

Ghiani, G., Laporte, G., & Musmanno, R. (2013). *Introduction to Logistics Systems Management* (2. ed.). Chichester: Wiley.

GMPL (2021). *GNU MathProg modeling language.* Retrieved February 2021, https://www.gnu.org/software/glpk/.

Google (2021a). *Distance Matrix API.* Retrieved February 2021, https://developers.google.com/maps/documentation/distance-matrix/overview.

Google (2021b). *Google OR-Tools.* Retrieved September 2021, https://developers.google.com/optimization.

Graphhopper (2021). *Products.* Retrieved February 2021, https://www.graphhopper.com/products/.

Grasas, A., & Ramalhinho, H. (2016). Teaching distribution planning: a problem-based learning approach. *The International Journal of Logistics Management, 27*(2), 377-394.

Guimarães, L., & Lima, R. (2021). Changes in teaching and learning practice in an undergraduate logistics and transportation course using problem-based learning. *Journal of University Teaching & Learning Practice, 18*(3), 11-27.

Hillier, F., & Lieberman, G. (2015). *Introduction to Operations Research* (10. ed.). New York et al.: McGraw-Hill.

Hung, W. (2011). Theory to reality: a few issues in implementing problem-based learning. *Education Tech Research Dev*(59), 529-552.

Inchainge (2021). *Discover The Fresh Connection*. Retrieved April 2021, https://inchainge.com/business-games/tfc/

Kanet, J. J., & Stößlein, M. (2008). Using a Supply Chain Game to Effect Problem-Based Learning in an Undergraduate Operations Management Program. *Decision Sciences Journal of Innovative Education, 6*(2), 287-295.

Kay, M. G. (2016). Matlog: Logistics Engineering using Matlab. *Journal of Engineering Sciences and Design, 4*(1), 15-20.

Mason, A. J. (2013). SolverStudio: A New Tool for Better Optimisation and Simulation Modelling in Excel. *INFORMS Transactions on Education*, 45-52.

Mattfeld, D., & Vahrenkamp, R. (2014). *Logistiknetzwerke: Modelle für Standortwahl und Tourenplanung* (2. ed.). Wiesbaden: Springer Gabler.

Microsoft (2021). *Distance Matrix API*. Retrieved February 2021, https://www.microsoft.com/en-us/maps/distance-matrix.

Neema, M., Maniruzzaman, K., & Ohgai, A. (2011). New Genetic Algorithms Based Approaches to Continuous p-Median Problem. *Networks and Spatial Economics*(11), 83–99.

OpenStreetMap. (2021a). Retrieved January 2021, https://www.openstreetmap.org/relation/8398124#map=16/40.7632/-73.9837.

OpenStreetMap (2021b). *OSRM API Documentation*. Retrieved February 2021, http://project-osrm.org/docs/v5.5.1/api.

Pyomo (2021). Retrieved Februray 2021, http://www.pyomo .org.

Rushton, A., Croucher, P., & Baker, P. (2017). *The Handbook of Logistics & Distribution Management* (6. ed.). London et al.: KoganPage.

Statistisches Bundesamt (2020). *Alle politisch selbständigen Gemeinden mit ausgewählten Merkmalen am 31.12.2020 (4. Quartal)*. Retrieved February 2021, https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/Administrativ/Archiv/GVAuszugQ/AuszugGV4QAktuell.html

Steglich, M., Feige, D., & Klaus, P. (2016). *Logistik-Entscheidungen: Modellbasierte Entscheidungsunter-stützung in der Logistik mit LogisticsLab* (2nd completely revised and expanded edition ed.). Berlin et al.: De Gruyter Studium.

Sweeney, D. C., Campbell, J. F., & Mundy, R. (2010). Teaching Supply Chain and Logistics Management Through Commercial Software. *The International Journal of Logistics Management, 21*(2), 293-308.

Toolbox (2013). Retrieved February 2021, https://tu-dresden.de/bu/wirtschaft/bwl/log/studium/materialien/toolbox.

Williams, H. (2013). *Model Building in Mathematical Programming* (5. ed.). Chichester: Wiley.

Winkels, H.-M. (2012). Modellbasiertes Logistikmanagement mit Excel: Lösungen von Problemen in der Logistik unter Verwendung der Tabellenkalkulation. Hamburg: DVV.

Flexible Prozessautomatisierung durch Integration der Prozess-Engine Camunda in eine generisch konfigurierbare Stamm-datenapplikation

Prof. Dr. Peter Edelmann

Prof. Dr. Armin Wagenknecht

M.Sc. Michel Jonathan Schmitz

Technische Hochschule Mittelhessen
Wilhelm-Leuschner-Straße 13
61169 Friedberg
Peter.Edelman@mnd.thm.de

Technische Hochschule Mittelhessen
Wiesenstraße 14
35396 Gießent
Armin.Wagen-knecht@mni.thm.de

Technische Hochschule Mittelhessen
Wiesenstraße 14
35396 Gießen
Michel.Jonathan.Schmitz@its.thm.de

ABSTRACT

Bei komplexen Verwaltungsprozessen ist die Vielfalt möglicher Prozessabläufe oft nicht vorhersehbar oder eine vollständige Modellierung nicht wirtschaftlich. WissensarbeiterInnen benötigen Möglichkeiten, um auf Abweichungen vom definierten Prozessablauf reagieren zu können. Oft werden daher neben dem Workflow-Management-System separate Applikationen zur Stammdatenverwaltung sowie zur Durchführung der nicht modellierten Prozessabläufe erstellt. Die vorliegende Arbeit untersucht die Hypothese, dass die Flexibilität für WissensarbeiterInnen durch Integration einer Prozess-Engine (Produkt Camunda) in eine generisch konfigurierbare NoSQL-Neo4J-basierte Stammdatenapplikation erhöht werden kann, durch Entwicklung eines Prototyps. In diesem wird die Prozess-Engine ausschließlich auf die Kontrolle des Workflows reduziert. Die Engine delegiert die Bearbeitung der im Prozess angefallenen Daten als BenutzerAufgabe an die Stammdatenapplikation und schränkt dabei den Aktionsraum der WissensarbeiterInnen durch Rechte ein. Die Integration bietet WissensarbeiterInnen einerseits die Unterstützung durch BPMN-definierte Workflows und andererseits die Flexibilität einer manuell bedienbaren Stammdatenapplikation. Durch das vorliegende Verfahren kann etwa ein nicht modellierter Change-Prozess von WissensarbeiterInnen manuell ausgeführt werden.

SCHLÜSSELWÖRTER

Prozessautomatisierung, Prozessflexibilisierung

EINLEITUNG

Die Prozessmodellierung kämpft mit dem Problem, die Balance zwischen Flexibilität und Strukturiertheit von Prozessabläufen zu finden.

Die Modellierung aller Ausführungsalternativen für die komplexe Prozesse eines Büroumfeldes in einem Workflow-Management-Systeme ist oft wegen des hohen Aufwands nicht sinnvoll.

WissensarbeiterInnen (Personen, deren Aufgabe es ist, Wissen zu erzeugen, zu verbreiten und anzuwenden (Davenport 2005, S. 10) benötigen vielfältige Möglichkeiten, um auf Abweichungen vom definierten Prozessablauf reagieren zu können.

Oft werden daher neben dem Workflow-Management-System separate Applikationen zur Stammdatenverwaltung sowie zur Durchführung der nicht modellierten Aktivitätsflüsse erstellt.

Die vorliegende Arbeit untersucht, ob die Flexibilität für WissensarbeiterInnen durch Integration eines Workflow-Management-Systems in eine generisch konfigurierbare Stammdatenapplikation erhöht werden kann.

Die Integration könnte einerseits die Unterstützung durch BPMN definierte Workflows bieten und andererseits die Flexibilität einer manuell bedienbaren Applikation, welche beliebige Aktivitätsflüsse innerhalb des durch das zugehörige Rechtesystem eingeschränkten Aktionsraums erlaubt.

STAND DER FORSCHUNG

Für die Modellierung von Prozessdefinitionen existieren zwei verschiedene Paradigmen.

Bei der imperativen Prozessmodellierung (Fahland et al. 2009, S. 2)) werden alle Ausführungsmöglichkeiten und Zustände des Models definiert. Im Falle einer neuen Alternative, muss diese explizit dem Modell hinzugefügt werden (Pesic 2008, S. 80). Fahland stellt die Frage, ob Workflow-Management-Systeme, die diese Art der imperativen Prozesse unterstützen, auch im Büroumfeld angewendet werden können und wirft dabei das Problem auf, dass Prozessbeteiligte im Büro mehr Möglichkeiten benötigen, um auf Abweichungen vom definierten Prozessablauf reagieren zu können (Reijers et al. 2003, S. 1-2).

Der Wunsch nach Flexibilität wird durch die deklarative Modellierung in den Vordergrund gerückt (van der Aalst et al. 2009, S. 99-100). Bei der deklarativen Prozessmodellierung werden Einschränkungen definiert, über die

implizit die möglichen Handlungsalternativen bestimmt werden. Das Hinzufügen von Einschränkungen bedeutet, dass Handlungsalternativen ausgeblendet werden (Pesic 2008, S. 80). Durch die wählbaren Handlungsalternativen können WissensarbeiterInnen den Prozess in einem gewissen Rahmen frei bearbeiten. Durch den Fokus auf Handlungsalternativen fehlt allerdings die Unterstützungsfunktion des Workflow-Management-Systems. Weniger gut ausgebildete Prozessbeteiligte können daher nicht effizient mit dem System arbeiten. Es entsteht dabei das Problem, die richtige Balance zwischen Flexibilität und Strukturiertheit von Prozessen zu finden (Ter Hofstede et al. 2009, S. 175-178; van der Aalst et al. 2009, S. 99-100).

Zu beiden Ansätzen existieren mittlerweile etablierte Standards der OMG. Die BPMN ermöglicht die imperative Prozessmodellierung (OMG, 2013) und Case Management Model and Notation (CMMN) die deklarative Prozessmodellierung (OMG 2016)

Bei dem Versuch eine Lösung für das Balance-Problem zu finden, gehen AutorInnen unterschiedliche Wege. Unter anderem ist die Hybridlösung BPMN-D entstanden. Die BPMN wird dabei um deklarative Elemente erweitert. Dazu wird ein Konvertierungsalgorithmus vorgestellt, der aus der deklarativen, einschränkungsbasierten Modellierungssprache Declare ein gültiges BPMN-D Modell erzeugt (De Giacomo et al. 2015, S. 88-99). Ein anderer Vorschlag, um die Vorteile der imperativen und deklarativen Modellierung zu nutzen, ist die Kombination der BPMN und CMMN (Hinkelmann and Pierfranceschi 2014, S. 10).

VORBEMERKUNG

Problemschilderungen, Konzepte, Szenarien und Prototyp orientieren sich zur besseren Verständlichkeit und im Hinblick auf eine gewünschte Praxisnähe an dem im Fachbereich MND der Technischen Hochschule Mittelhessen betriebenen Onlinesystem für den Verwaltungsprozess zur Abwicklung einer Abschlussarbeit, z.B. Masterarbeit, im Studiengang Wirtschaftsinformatik.

PROBLEME HERKÖMMLICHER PROZESSAUTOMATISIERUNGSSYSTEME

Geringe Flexibilität bei der Bearbeitung von HumanTasks

Herkömmliche Prozessautomatisierungen mit weit verbreiteten Prozess-Engines, z.B. dem Produkt Camunda, gewähren der AnwenderIn bei Ausführung einer Benutzeraufgabe (HumanTask) nur eine geringe Flexibilität.

So kann zum Beispiel der Prüfungsausschussvorsitz im Rahmen der Genehmigung einer Masterarbeit nur Thema und Stammdaten der StudentIn einsehen. Die besuchten Wahlpflichtfächer, die Themen vorhergehender Seminare, Praktika, oder Informationen über die betreuende Firma sind nicht verfügbar.

Möchte die ProzessadministratorIn dem Prüfungsausschussvorsitz für die betrachtete Benutzeraufgabe eine größere Flexibilität einräumen, so muss der Prozess mit

verschiedenen AutomaticTasks zur Datenbeschaffung und weiterer Webseiten (HumanTaskViews) zur Datendarstellung erweitert werden. Eine solche Erweiterung kann in der Regel nur von ProgrammierInnen mit Erfahrungen in den Bereichen BPMN, Webseiten, und Datenbanken vorgenommen werden.

Separate Applikationen zur Durchführung der im Prozess nicht modellierten Aktivitätsflüsse

Prozesse in der realen Bürowelt sind häufig sehr komplex. Bedingt durch diese hohe Komplexität werden in der Praxis oft nicht alle erforderlichen Datenmanipulationen und Entscheidungsmöglichkeiten durch zugehörige Pfade der Prozesse bzw. Subprozesse abgebildet. Im betrachteten Beispiel der Masterarbeit mag etwa kein Change-Prozess integriert sein, der eine nachträgliche Änderung des Themas oder eine Verschiebung des Abgabedatums wegen Krankheit abbildet.

Auch ein Bedienfehler, beispielsweise das Absenden fehlerhafter Eingaben, kann oft nicht kompensiert werden. Die im Rahmen der HumanTask von der Prozess-Engine zur Verfügung gestellten Webseiten (HumanTaskViews) bilden die einzige Möglichkeit für die BenutzerIn, Daten anzusehen und zu editieren.

Daher existieren oft neben dem System zur Prozessautomatisierung ein oder mehrere Applikationen zur Verwaltung der im Rahmen des Prozesses anfallenden Stammdaten. Diese Applikationen erlauben es den WissensarbeiterInnen, nicht modellierte Aktivitätsflüsse im Rahmen der ihr erteilten Rechte und basierend auf ihrem Prozessverständnis manuell auszuführen.

Dazu gehören etwa eine Korrektur der Stammdaten außerhalb des Prozesses, eine Bearbeitung der Stammdaten im Falle eines Prozessabbruchs oder eines nicht modellierten Prozesspfades.

FLEXIBLE PROZESSAUTOMATISIERUNG DURCH INTEGRATION DER PROZESS-ENGINE CAMUNDA IN EINE GENERISCH KONFIGURIERBARE STAMMDATEN-APPLIKATION

Der vorliegende Ansatz zielt darauf ab, die oben genannten Probleme durch eine Kopplung der Prozess-Engine mit einer generisch konfigurierbaren Stammdatenapplikation zu reduzieren.

Dabei übernimmt jede der beiden Komponenten, Prozess-Engine und Stammdatenapplikation genau die Aufgaben für die sie besonders geeignet ist und konstruiert wurde. D.h. die Stammdatenapplikation beschafft und präsentiert der AnwenderIn im Rahmen einer HumanTask die gewünschten Daten und die Prozess-Engine kontrolliert den Workflow (vgl. Abbildung 1).

Die imperative BPMN-Beschreibung des durch die Prozess-Engine abzuarbeitenden Workflows wird mit der

deklarativen Beschreibung des Aktionsraumes einer AnwenderIn durch Rollen und Rechte in der Stammdatenapplikation kombiniert.

Die Webseite kann durch die Verzahnung mit der Prozess-Engine nur im Rahmen der HumanTask verwendet werden, nicht aber, um nachträglich z.B. das Thema oder
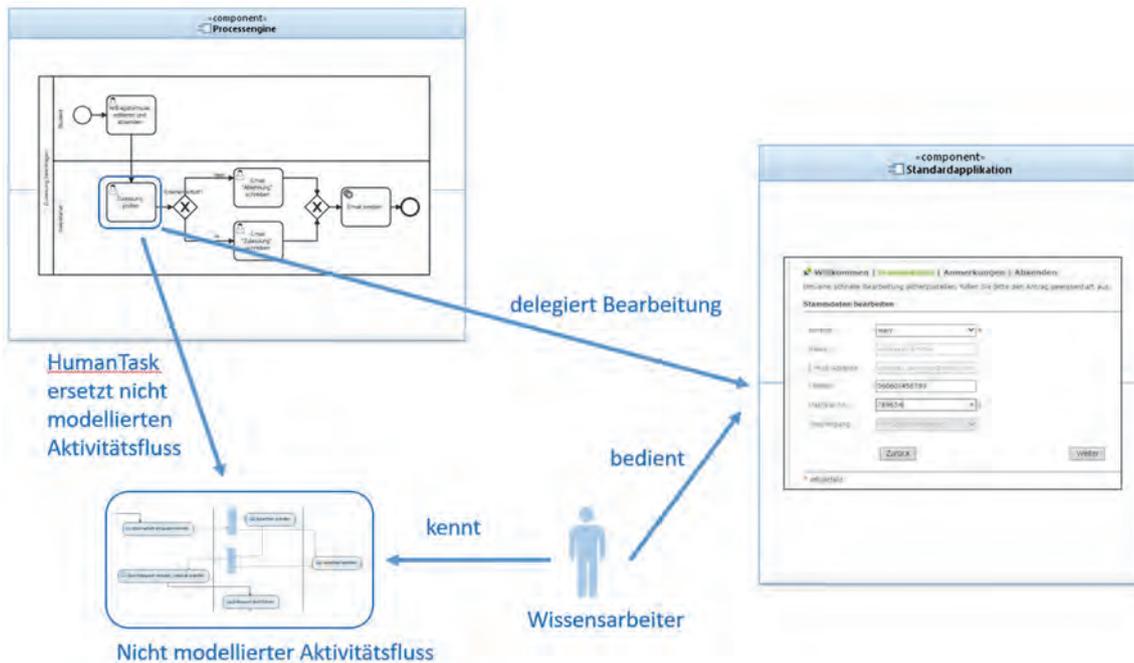
Abbildung 1: Delegation Aktivitätsfluss an Wissensarbeiter mittels HumanTask

FACHLICHE ARCHITEKTUR DES PROTOTYPS

Reduktion der Prozess-Engine auf die Workflow

Steuerung

Das Kommunikationsdiagramm in Abbildung 2 zeigt Objekte und Methodenaufrufe zur Darstellung der Kollaboration für eine herkömmliche Verwendung einer Prozess-Engine (grüne Objekte sind der Domäne Stammdatenverwaltung, blaue Objekte sind der Domäne Prozess-Engine zuzuordnen).
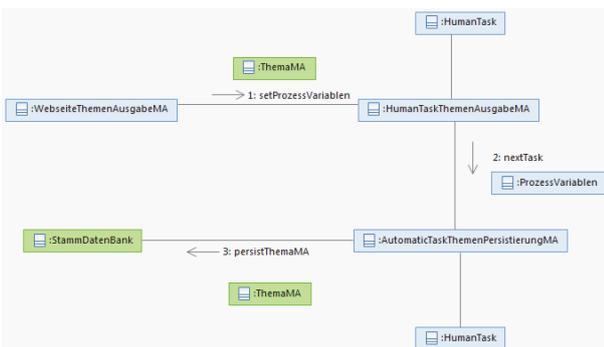
Abbildung 2: Kopplung der Web-Seiten und des Datentransports an die Prozesssteuerung

Die Angabe des Themas der Masterarbeit erfolgt im HumanTaskView also in der der HumanTask zugeordneten Webseite. Das Thema wird in die Prozessvariablen der Prozess-Engine transportiert und in der nachfolgenden AutomaticTask in die Stammdatenbank persistiert.

den Abgabezeitpunkt zu korrigieren.
Nach einem Abbruch des Prozesses kann aufgrund der Kopplung mit der Prozess-Engine keine den Human-Tasks zugeordnete Webseite bedient werden.

Abbildung 3 zeigt die Entkopplung der Webseiten und des Datentransports von der Prozesssteuerung (vgl. Abbildung 2 zur Erläuterung des Kommunikationsdiagramms).
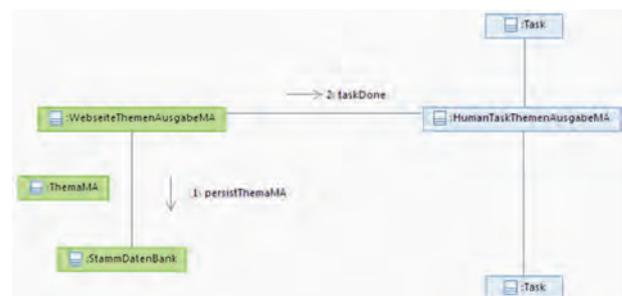
Abbildung 3: Entkopplung der Web-Seiten und des Datentransports von der Prozesssteuerung

Die Angabe des Themas der Masterarbeit und die Persistierung in die Stammdatenbank erfolgen durch Bedienen der Stammdatenapplikation.
Die Aufgabe der angekoppelten Prozess-Engine reduziert sich auf die Steuerung der Aktivitäten der im Prozess Masterarbeit beteiligten AkteurInnen.
So kann zum Beispiel eine Änderung des Themas oder Abgabezeitpunktes ohne Beteiligung der Engine direkt in der Stammdatenapplikation erfolgen. Ggf. kann die AkteurIn innerhalb des durch das Rechtesystem einge-

schränkten Aktionsraums einen nicht modellierten Aktionsfluss manuell ohne Engine ausführen. Nach einem Ausfall der Prozess-Engine kann der Prozess ebenfalls manuell ohne Engine weitergeführt werden.

Abbildung 4 zeigt den Unterschied zwischen den Mechanismen einer herkömmlichen Prozess-Engine und dem vorliegenden Ansatz basierend auf der losen Kopplung im direkten Vergleich (vgl. Abbildung 2 zur Erläuterung der Kommunikationsdiagramme).
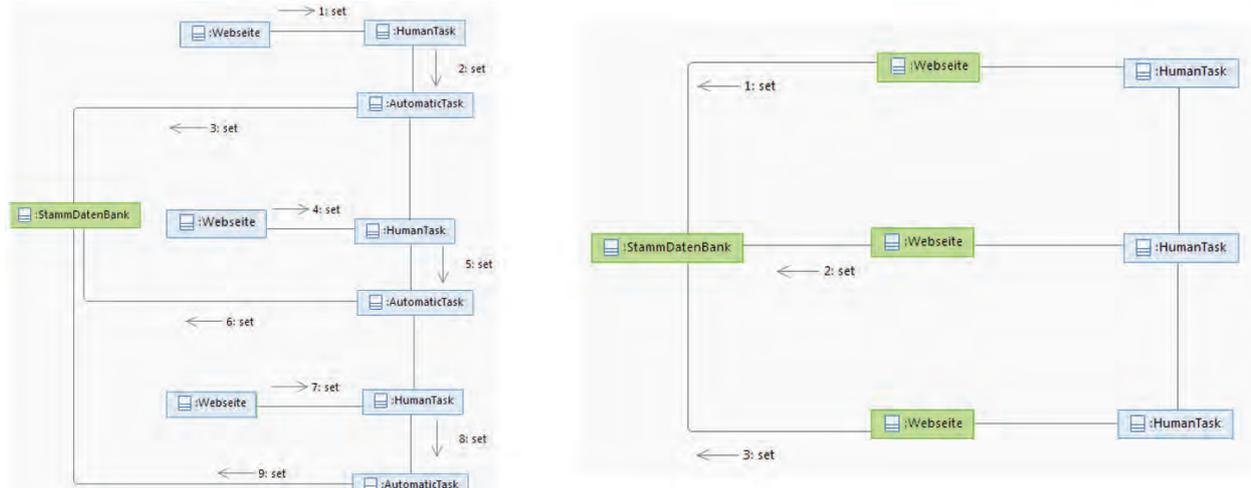
Abbildung 4: Unterschied Mechanismen bei enger Kopplung (links) und loser Kopplung (rechts)

Die Webseiten sind bei loser Kopplung vollständig in der manuell von der WissensarbeiterIn bedienbaren Stammdatenapplikation integriert und lediglich für die Workflow-Steuerung mit der Prozess-Engine gekoppelt.

Generisch konfigurierbare Stammdatenapplikation basierend auf NoSQL-Metadaten

Im Hinblick auf eine verbesserte Flexibilität für die ProzessadministratorIn wurde die Prozess-Engine an eine generisch konfigurierbare Stammdatenapplikation gekoppelt. Die Stammdatenapplikation unterstützt Listen- und Detailansichten zum Erzeugen, Lesen, Ändern und Löschen (CRUD) von generisch konfigurierbaren Entitäten. Die Zugriffssteuerung erfolgt einheitlich über Rollen, Rechte und Filter. Diese spezifizieren, die für die AnwenderIn sichtbaren Entitäten und Navigationsmöglichkeiten. Basis der Applikation bildet eine Neo4j-NoSQL-Datenbank. Diese unterstützt die Abspeicherung von Stammdaten in Form von gelabelten Nodes und Relations. Sie liefert Attribute und Relations in Form eines Arrays von String basierten Key-Value Pärchen. Die zur Laufzeit erzeugten Ansichten basieren auf Metadaten zu den Entitätstypen. Die Einschränkung der AnwenderIn und Validierung der eingegebenen Stammdaten erfolgt durch die Definition von MetaNodes. Diese legen für einen Entitätstyp die Bezeichnungen, Datentypen und Werteinschränkungen der Attribute sowie der Relationen zu anderen Entitätstypen fest. Die Modellie-

rung von Rechten und Filtern in den Metadaten ermöglicht die Einschränkung des Zugriffs mittels der Neo4j Cypher Query Language basierend auf Graph Patterns bestehend aus Anker und Pfad. Die generische Stammdatenapplikation unterstützt eine anwendungsspezifische Codierung über konfigurierbare „SpecialServices". Sie erlaubt z.B. das „Claimen" („Claimen" bezeichnet die Aktion, das eine NutzerIn eine Aufgabe für sich zur Bearbeitung beansprucht) einer HumanTask, ListItems mit speziellen Informationen anzureichern, allen Personen

einer Liste Emails zu senden oder die Generierung eines spezifischen Reports.
Dazu werden im Meta-Modell alle anwendungsspezifisch zu codierenden Aktionen durch MetaActions modelliert und durch verschiedene Relations, z.B. forwardToViewAction, forwardToServiceAction, onCreateAction, mit der betroffenen MetaNode verbunden.
Die MetaAction spezifiziert den SpecialService bzw den SpecialView. Die GUI-Handler delegieren Requests für MetaNodes ggf. an zugeordnete SpecialServices.

Repräsentanten der prozessrelevanten Informationen in der Stammdatenapplikation

Im Falle einer Prozessbearbeitung verwaltet die Prozess-Engine meist eine sogenannte Prozessvariable, welche alle prozessrelevanten fachlichen Informationen kapselt.
Da diese Informationen in der vorliegenden Arbeit als Neo4j-Nodes in der Stammdatenapplikation verwaltet werden, wird zu jedem Prozess eine Metadaten-Struktur editiert, welche die prozessrelevanten fachlichen Informationen in Form von Entitäten repräsentiert.
Diese Datenstruktur beinhaltet immer eine ProcessAggregat-Node, welche alle prozessrelevanten Entitäten aggregiert. Im Beispiel des Prozesses zur Verwaltung der Masterarbeit werden etwa Nodes zu dem Studierenden, der Firma, der BetreuerIn und zur Spezifikation von Thema und Bearbeitungszeitraum der Masterarbeit sowie eine aggregierende MasterArbeitProcessAggregat-Node erzeugt (vgl. Abbildung 5).
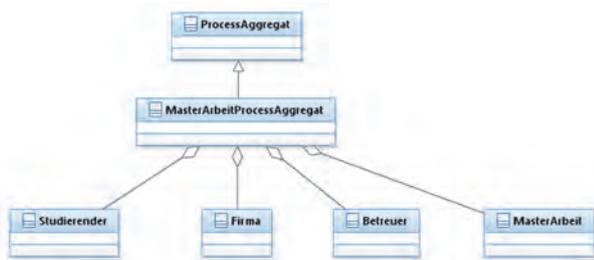
Abbildung 5: Repräsentanten der fachlichen Prozessinformationen

Repräsentanten des Prozesses und der HumanTasks in der Stammdatenapplikation

Damit eine AkteurIn eine HumanTask im Rahmen der Stammdatenapplikation „claimen" kann, wird ein Repräsentant des von der Prozess-Engine erzeugten internen Camunda Prozesses und der zugehörigen zu bearbeitenden HumanTask in der Stammdatenapplikation benötigt. Daher wurden die entsprechenden Entitäten EngineProzess-Node und EngineHumanTask-Node in der Metadaten Struktur spezifiziert.

Beide Entitäten repräsentieren lediglich Elemente der Camunda Prozess-Engine und sind zu unterscheiden von den Entitäten, welche prozessrelevante fachliche Informationen kapseln.

Im Beispiel des Prozesses zur Verwaltung der Masterarbeit wird für den gestarteten Prozess eine EngineProzess-Node und für die auszuführende HumanTask „Masterarbeit Themenausgabe" eine EngineHumanTask-Node instanziert (vgl. Abbildung 6).

Die EngineProzess-Node und EngineHumanTask-Node enthalten die entsprechenden IDs der internen Elemente der Camunda-Prozess-Engine.

Die EngineProzess-Node aggregiert die zugehörigen EngineHumanTask-Nodes.

Die EngineProzess-Node referenziert die zugehörige ProcessAggregat-Node. Diese Referenz repräsentiert die Verbindung von der Prozesswelt der Engine zur Stammdatenwelt.
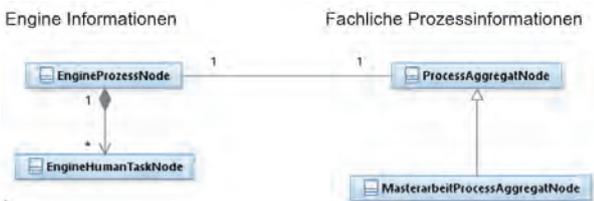
Abbildung 6: Prozess-Repräsentanten

Die Camunda-Prozess-Engine referenziert die EngineProzess-Node mittels Neo4J-ID im Rahmen der Prozessvariablen und das Interface zur Ansprache der Stammdatenapplikation über eine URL (REST Calls).

Rechte und Rollen

Task-Rollen sind temporäre Rollen, die nur zu dem Zeitpunkt aktiviert werden, wenn die BenutzerIn eine HumanTask bearbeitet. Der Name einer Task-Rolle ist per

Konvention identisch mit der Element-ID des korrespondierenden BPMN-Elements. Die Rolle beschreibt den Aktionsraum, den eine BenutzerIn im Rahmen der HumanTask verwenden können. Die BenutzerIn kann die Aufgabe innerhalb des vorgegebenen Aktionsraums, bearbeiten und ist dabei nicht auf feste Reihenfolgen von Ansichten angewiesen. Sie ist jedoch auf die für ihre Aufgabe relevanten Daten eingeschränkt. BenutzerInnen, die eine Lane-Rolle besitzen, sind autorisiert alle Aufgaben einer BPMN-Lane eines Prozesses zu bearbeiten. Eine Lane-Rolle kann Personengruppen zugeordnet werden. Die Zuordnung zu den Aufgaben der Lane erfolgt über einen Filter, der alle Benutzer-Aufgaben einer BPMN-Lane basierend auf Benennungskonventionen auffindet.

Einer eingeloggten AnwenderIn (UserIn) ist eine Role-Node zugeordnet und dieser wiederum mehrere Right-Nodes. Jede Right-Node aggregiert eine CRUDE-Specification für ein Node Label und Filter zur Selektion der Nodes des angegebenen Labels. CRUDE steht für create, read, update, delete, execute. Die Abbildung 7 zeigt die zugehörige Metadaten-Struktur.
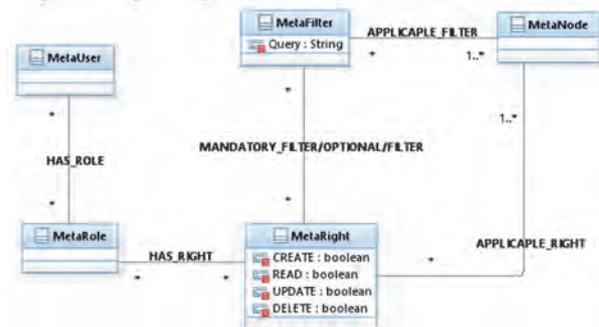
Abbildung 7: Rechte, Rollen und Filter

Filter werden mittels Neo4J-Cypher-Queries spezifiziert. Zu diesem Zweck muss der Startknoten der Query in Form der AnchorNode spezifiziert werden.

Durch die Abhängigkeit der Abfrage von der eingesetzten AnchorNode entstehen verschiedene Ergebnismengen.

Für die Stammdatenapplikation wird in der Regel die eingeloggte AnwenderIn als Startknoten verwendet.

Da verschiedene AnwenderInnen auf dieselben Rollen verweisen, Rollen also nicht User-spezifisch definiert sind, muss die AnchorNode User-spezifisch gesetzt werden.

Daher ist die AnchorNode eine Association-Klasse für die Association-User zur Rolle vgl. Abbildung 8.
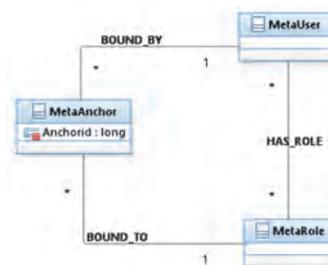
Abbildung 8: AnchorNode Metadaten-Struktur

Die AnchorNode kann erst zur Laufzeit gesetzt werden. Per Default wird für die Stammdatenapplikation die AnchorNode beim Einloggen vom Login-Controller im Rahmen der Rechte Zuweisung auf die Node-ID der eingeloggten AnwenderIn gesetzt. Im Falle einer Prozessbearbeitung wird in der Regel die ProcessAggregat-Node als AnchorNode verwendet (vgl. Abbildung 9). Bei Annahme einer HumanTask wird die AnchorNode von dem der Aktion ClaimTask zugeordneten Service im Rahmen der Rechtezuweisung auf die ProcessAggregat-Node des Prozesses gesetzt.

Für die Zuweisung einer Rolle zu einer AnwenderIn werden die Relationen hasRole, hasTempRole, hasInactiveRole, unterschieden.

Zur temporären Rechtevergabe an eine AnwenderIn als

Die Prozess-Engine führt die HumanTask zur Spezifikation der Masterarbeit aus und beauftragt auf diese Weise die BetreuerIn mit der Spezifikation der Masterarbeit. Die BetreuerIn akzeptiert die Aufgabe und spezifiziert unter anderem die Annahme der Betreuung, Thema und Abgabedatum (vgl. Szenario „Claimen einer HumanTask über Repräsentanten").

Die BetreuerIn quittiert die Erledigung der Aufgabe (vgl. Szenario „Beenden einer HumanTask").

Die Prozess-Engine holt aus der Datenbank der Stammdatenapplikation die Spezifikation der Masterarbeit und führt das Gateway zur Entscheidung über die Annahme der Betreuung aus (vgl. Szenario „Entscheidungsfindung an einem Gateway"). Die Prozess-Engine sendet eine Mail mit der Spezifikation der Masterarbeit an die

Abbildung 9: AnchorNode-Beispiel

AkteurIn einer HumanTask „hängt" der der Aktion ClaimTask zugeordnete Service der Stammdatenapplikation die zugeordnete Rolle in die Liste der aktiven temporären Rollen der AnwenderIn.

Dieser Service befüllt ferner die der Rolle und der eingeloggten AnwenderIn zugeordnete Anchor-Node mit der ID der Process Aggregat-Node.

Zum Rechteentzug entfernt der der Aktion TaskDone zugeordnete Service die der HumanTask zugeordnete Rolle wieder aus der Liste der aktiven temporären Rollen und „hängt" sie in die Liste der nicht aktiven Rollen der eingeloggten AnwenderIn. Die temporäre Rolle wird priorisiert und ersetzt alle anderen Rollen der eingeloggten AnwenderIn.

Die AkteurIn hat während der Erledigung der Task daher nur die der Task zugeordnete temporäre Rolle und ist somit eingeschränkt auf die zur Erledigung der HumanTask benötigten Rechte.

Zerlegung des Prozesses in Use-Case-Szenarien

Der sehr vereinfacht dargestellte Verwaltungsprozess zur Abwicklung einer Masterarbeit beinhaltet die folgenden Use-Case-Szenarien (siehe Abbildung 10 auf der nächsten Seite).

Die StudentIn füllt im Rahmen der Stammdatenapplikation einen Antrag zur Erstellung der Masterarbeit aus und startet damit in der Prozess-Engine den Prozess.

StudentIn. Schließlich schließt die Prozess-Engine den Prozess ab. Einige dieser Szenarien werden in den folgenden Kapiteln ausführlicher erläutert.

Szenario „Claimen einer HumanTask über "Repräsentanten"

Abbildung 11 (auf der übernächsten Seite) zeigt die Erzeugung eines Repräsentanten der HumanTask der Prozessengine und das darauffolgende Öffnen einer der zugehörigen Webseiten zur Task-Bearbeitung in der Stammdatenapplikation. Zugriffe auf Services sind nicht dargestellt (vgl. Abbildung 2 zur Erläuterung des Kommunikationsdiagramms).

Die Prozess-Engine Camunda bietet die Möglichkeit eigene Ereignis-Listener-Klassen zu erstellen. Dabei handelt es sich um Java-Klassen, die ausgeführt werden, wenn das Token der Prozessinstanz das BPMN-Element erreicht und das konfigurierte Ereignis ausgelöst wird.

Die von der Prozess-Engine aktivierte HumanTask delegiert innerhalb eines solchen Ereignis Listeners die Ausführung der Task an die Stammdatenapplikation.

Zu diesem Zweck erzeugt die Prozess-Engine im Verlauf des Prozesses für jede von einer AkteurIn auszuführende HumanTask über den ProcessAndTaskNotificationService der Stammdatenapplikation eine EngineHumanTask-Node als zugehörigen Repräsentanten innerhalb der Stammdatenapplikation, so z.B. auch für die HumanTask „Masterarbeit Themenausgabe".

innerhalb der Komponente Prozessverwaltung der Stammdatenapplikation ausgeführt. Die Pflege eines Rechtesystems in der Prozess-Engine entfällt.

Die Liste der EngineHumanTask Nodes wird dabei durch Filterung (Cypher Query) auf diejenigen Nodes reduziert, welche mindestens eine indirekte Beziehung
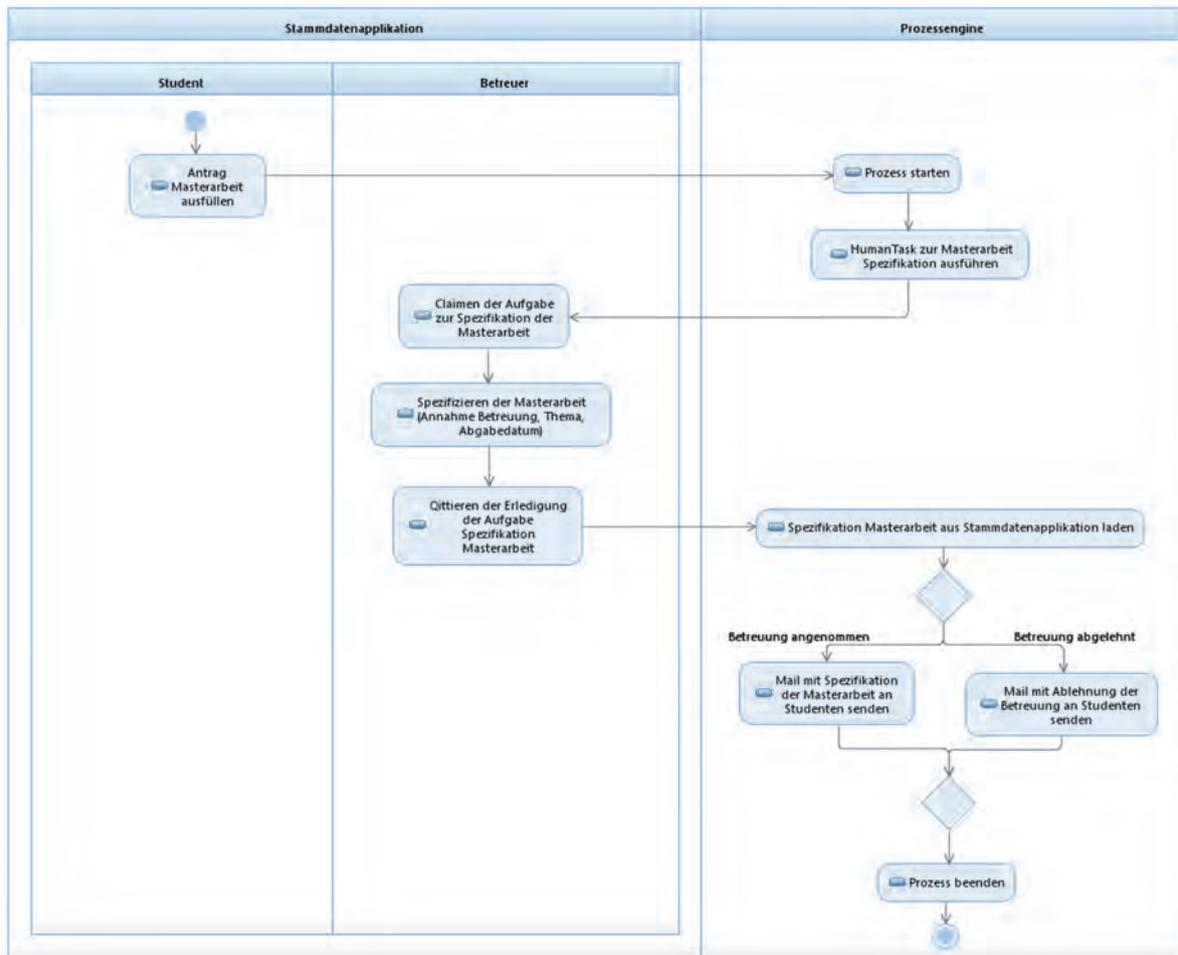
Abbildung 10: Aktivitätsdiagram Student-Betreuer-Engine

Der aufgerufene ProcessAndTaskNotificationService verbindet diese EngineHumanTask-Node mit derjenigen Node, welche die der HumanTask zugeordnete Rolle repräsentiert (Das Node Label der Rolle entspricht per Konvention dem Namen der HumanTask). Eine entsprechende Zuordnung erfolgt innerhalb der Prozessvariablen von Camunda.

Die innerhalb der Camunda Prozess-Engine vergebene ID der erzeugten Instanz der HumanTask wird dabei als Referenz in ein Attribut der repräsentierenden EngineHumanTask-Node gesetzt.

Schließlich wird die EngineHumanTask-Node als offene Aufgabe der bereits existierenden EngineProzess-Node hinzugefügt.

BenutzerInnen können Aufgaben in Form von EngineHumanTask-Listen und Detail-Seiten einsehen. Das Beanspruchen („Claimen") und Freigeben einer Aufgabe wird nicht an die Prozess-Engine kommuniziert, sondern

zu einer der inaktiven Rollen der eingeloggten AnwenderIn haben.

Das „execute" Recht ermöglicht der AkteuerIn das „Claimen" der selektierten EngineHumanTask.

Durch das „Claimen" weist die Stammdatenapplikation im Rahmen eines ClaimTaskService der AkteurIn temporär die der auszuführenden HumanTask zugeordnete Rolle zu. Dies geschieht durch „einhängen" der Rolle in die Liste der aktiven temporären Rollen der AnwenderIn. So erhält die AkteurIn den Zugriff auf alle Entitäten gemäß der durch die Rolle spezifizierten Rechte.

Im Beispiel erhält die betreuende Person der Masterarbeit etwa das Recht, die Nodes StudentIn und Firma zu lesen, die Betreuer-Node zu editieren sowie eine Node MasterArbeit zu erzeugen.

Der Prüfungsausschussvorsitz erhält schließlich die Möglichkeit, die MasterArbeit-Node zu editieren um Fristen und Thema zu prüfen und ggf. ändern zu können.

Darüber hinaus kann der Vorsitz die Nodes StudentIn, Firma und BetreuerIn lesen. Das Lesen der Betreuer-Node ermöglicht etwa die Prüfung der Qualifikation der FirmenbetreuerIn.

Szenario „Entscheidungsfindung an einem Gateway"

Eine AutomaticTask transportiert die im Gateway zur Entscheidung benötigten Daten mittels generischem DataAccessService aus der Stammdatenapplikation in
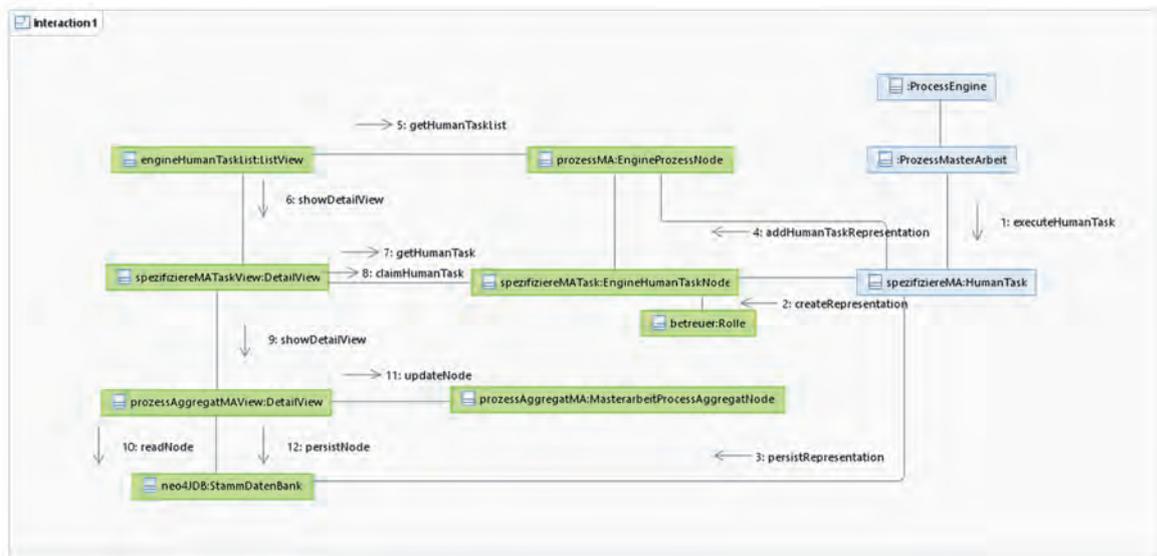
Abbildung 11: Erzeugung eines HumanTask Repräsentanten und „Claimen"

Ggf. könnte man seine Rechte noch dahingehend erweitern, dass er z.B. die von der StudentIn besuchten Wahlpflichtfächer, die Themen vorhergehender Seminare und Praktika einsehen kann.

Aufgrund der Priorisierung der temporären Rollen gegenüber den dauerhaften Rollen wird der Aktionsraum und somit auch die Ansicht der BenutzerIn auf die Bearbeitung der Aufgabe reduziert.

Die BenutzerIn kann die Aufgabe innerhalb des vorgegebenen Aktionsraums, bearbeiten und ist dabei nicht auf feste Reihenfolgen von Ansichten angewiesen.

Szenario „Beenden einer HumanTask"

Wenn die AkteurIn die Erledigung der HumanTask mittels des Buttons TaskDone im ListView der EngineHumanTask-Nodes quittiert, so meldet der zugehörige TaskDoneService die zur Beendigung der HumanTask benötigten Informationen zur Identifikation von Prozess und HumanTask an die Prozess-Engine.

Fachliche Daten, im Beispiel etwa die Spezifikation der Masterarbeit, wurden bereits direkt an die Datenbank der Stammdatenapplikation übermittelt und werden daher nicht an die Camunda-Prozess-Engine übertragen.

Der TaskDoneService entfernt die der HumanTask zugeordnete Rolle wieder aus der Liste der aktiven temporären Rollen und hängt sie in die Liste der nicht aktiven Rollen der eingeloggten AnwenderIn.

Der ordnungsgemäße Abschluss der HumanTask wird zusammen mit einem Zeitstempel in der EngineHumanTask Node eingetragen und diese aus der Liste der offenen Aufgaben entfernt. Zum Einsehen der Historie wird sie in der Stammdatenapplikation weiter verwaltet.

die Prozessvariablen der Prozess-Engine. Die Spezifikation der benötigten Informationen erfolgt über Cypher Query Strings.

Der von der Prozess-Engine zur Entscheidung benötigte Expression Language Ausdruck wird mittels der Prozessvariablen im Rahmen der Prozessmodellierung formuliert und von der Prozessengine zur Laufzeit ausgewertet.

Im Beispiel ergibt sich etwa aus der Spezifikation der Masterarbeit, dass die BetreuerIn die Arbeit angenommen hat.

FAZIT

Die Prozessmodellierung kämpft mit dem Problem, die Balance zwischen Flexibilität und Strukturiertheit von Prozessabläufen zu finden.

In dieser Arbeit wurde ein imperativer Ansatz als Basis verwendet. Durch die feste mittels BPMN modellierte Ablaufsteuerung können BenutzerInnen eindeutig geführt werden. Sobald die BenutzerIn jedoch eine Aufgabe beansprucht erhält sie durch die zugeordnete Rolle die für eine WissensarbeiterIn wünschenswerte Flexibilität zur Bearbeitung der Aufgabe.

Dies wird erreicht durch eine Reduktion der Prozess-Engine auf die reine Workflow-Steuerung und eine rollenbasierte Delegation der Benutzer-Aufgaben an eine generisch konfigurierbare Stammdatenapplikation.

Aus dem betrachteten Verfahren ergeben sich die nachfolgenden Vorteile:

Erhöhte Flexibilität für WissensarbeiterInnen

Zur Bearbeitung einer HumanTask können der AkteurIn über die zugeordnete Rolle alle Sichten und Funktionen zur Verfügung gestellt werden, welche die Stammdatenapplikation bietet.

Die BenutzerIn kann frei entscheiden, welche Sicht sie wählt oder welche Aktionen sie innerhalb der Bearbeitung ausführt.

Auf diese Weise kann ein Aktivitätsfluss, welcher zur Modellierungszeit nicht vollständig analysiert oder wegen zu hoher Komplexität nicht modelliert wurde, an eine WissensarbeiterIn delegiert werden. Diese kann den Aktivitätsfluss im Rahmen der ihr erteilten Rechte und basierend auf ihrem Prozessverständnis manuell ausführen. Ein an eine WissensarbeiterIn delegierter nicht in der Prozess-Engine modellierter Aktivitätsfluss birgt die Gefahr eines nicht rechtskonformen Ablaufs und kann derzeit nur durch die Einschränkung der Rechte begrenzt werden. Der erhöhte Freiheitsgrad bzw. die Einschränkung der Rechte spiegelt das Vertrauen in die beteiligte WissensarbeiterIn wider. (vgl. auch weiter unten „Bedingte Freischaltung von Webseiten")

Manuelle Bearbeitung der fachlichen Prozessdaten

Die Stammdatenapplikation ermöglicht eine manuelle nicht von der Prozess-Engine getriggerte Bearbeitung der im Prozess angefallenen Daten. Auf diese Weise können z.B. Korrekturen an den im Rahmen einer HumanTask eingegebenen Daten vorgenommen werden. So kann z.B. das Thema einer Masterarbeit oder das Abgabedatum geändert werden. Ebenso kann ein Prozess für einen nicht modellierten Aktivitätsfluss manuell beendet werden.

Vereinfachtes Verfahren zur Erstellung von automatisierten Prozessen

Prozesse können alleine durch Konfiguration der Entitäten einer Domäne in der Stammdatenapplikation mittels MetaNodes und der Modellierung des BPMN-Workflows in der Prozess-Engine auf hoher Abstraktionsebene ablauffähig spezifiziert werden.

Ein Codieren von Webseiten für HumanTasks entfällt, da die Listen- und Detailsichten zur Datenein- und -ausgabe ebenso wie Sichten zur Vernetzung der Daten auf Basis der spezifizierten MetaNodes erzeugt werden.

Ein aufwändiges Codieren der HumanTask-Prozess-Elemente, z.B. mittels Groovy oder Java, entfällt, da kein Datentransport von den Webseiten in die Prozessvariablen und umgekehrt organisiert werden muss.

Ebenso entfällt ein Transport der fachlichen Daten von Prozessvariablen der Prozess-Engine in die Stammdatenapplikation und umgekehrt.

Daten müssen lediglich für Gateways aus der Stammdatenapplikation in die Prozessvariablen der Engine transportiert werden.

Separation des Workflows vom fachlichen Datenfluss

Das betrachteten Verfahren unterstützt die fachliche Separation des Workflows von den im Rahmen der HumanTasks durchzuführenden Aktionen und ermöglicht damit eine den verschiedenen Abstraktionsebenen angemessene Modellierung und Kapselung in verschiedene Komponenten.

So tangiert z.B. eine Hinzunahme eines Attributes in einer im Rahmen einer HumanTask zu editierenden Entität nicht den modellierten BPMN-Prozess.

Z.B. könnte lediglich durch Editieren der die Entität Masterarbeit spezifizierenden MetaNode ein Attribut Seitenzahl hinzugefügt werden. Der Workflow ist nicht betroffen und die der HumanTask zur Ausgabe des Themas zugeordnete Webseite wird zur Laufzeit aus den Metadaten generiert.

Durch die REST-basierte Kommunikation der Prozess ausführenden Prozess-Engine mit der Stammdatenapplikation können für beide Komponenten unterschiedliche Technologien verwendet werden. Auf diese Weise kann die Workflow Steuerung von der Stammdatenpflege auch technologisch entkoppelt werden.

Für das betrachtete Verfahren ergibt sich aber auch noch Verbesserungspotential bezogen auf nachfolgende Aspekte.

Codeanpassungen der Stammdatenapplikation zur Kopplung an die Prozess-Engine

Die Einbindung einer Prozess-Engine erfordert Codeanpassungen der Stammdatenapplikation z.B. in Form einer Ergänzung der Domäne zur Prozess- und Taskverwaltung sowie der Zuordnung von Prozessen zu Entitäten anderer Domänen. Dies gelingt bei der betrachteten generischen Stammdatenapplikation durch ein Konfigurieren von Entitäten und der Hinzunahme von „SpecialServices" für das Starten von Prozessen und das „Claimen" von HumanTasks. Für existierende nicht generische Stammdatenapplikationen ergeben sich an dieser Stelle aufwändige Modifikationen, die nur partiell durch die Wiederverwendung der existierenden Webseiten der Stammdatenapplikation für die Prozessautomatisierung kompensiert werden.

Austauschbarkeit der Prozess-Engine nur eingeschränkt möglich

Das Interface zur Kopplung der Stammdatenapplikation mit der Prozess-Engine bedient die bisher verwendete Camunda-API. Eine andere Prozess-Engine muss ähnliche Konzepte aufweisen, um sich für diese Verwendung zu qualifizieren.

Ein Austausch wird durch die minimale Nutzung des Funktionsumfangs der Camunda-Prozess-Engine für die reine Workflowsteuerung und die Anbindung mittels einfacher REST-Schnittstelle unterstützt.

Erschwert wird ein Austausch im Falle einer Prozess-Engine, die BPMN-Elementen eine andere Semantik zuordnet oder sogar eine andere Prozessbeschreibungssprache (z.B. Petri-Netze) verwendet.

Unvollständiges Errorhandling

Die aktuelle Implementation des evaluierten Beispielprozesses unterstützt keine Rückabwicklung durch Kompensation. Die generische Stammdatenapplikation unterstützt keine Datenbanktransaktionen. Systemreaktionen

für seltene Bedienfehler oder technische Defekte wurden nicht implementiert.

Evaluation komplexer Prozesse

Bisher wurde das System nur anhand eines eher einfachen Prozesses zur Abwicklung einer studentischen Abschlussarbeit getestet.

Dazu wurde im Labor für Softwaretechnologie ein Prototyp im Hinblick auf die beschriebenen wesentlichen Konzepte entwickelt und getestet. Eine weitergehende Evaluation durch Unterstützung mehrerer komplexer Prozesse im Fachbereich ist in Planung.

Bedingte Freischaltung von Webseiten durch Kombination der temporären Rollenvergabe mit einer statusab-hängigen Rollenvergabe

Eine bedingte Freischaltung von Aktionsmöglichkeiten (ähnlich dem CMMN-Wächter-Konzept) könnte in einer zukünftigen Version durch eine statusabhängige Rollenvergabe unterstützt werden.

Pflichten für den Aktor einer HumanTask

Zur Sicherstellung, dass eine „geclaimte" HumanTask nicht als „done" beendet wird ohne die eigentliche Aufgabe erfüllt zu haben, könnten der HumanTask im Rahmen der zugehörigen Rolle neben Rechten auch Pflichten zugeordnet werden. Eine solche Pflicht kann z.B. die Erzeugung einer Entität mit Pflichtfeldern sein, z.B. eine Masterarbeit-Node zur Kapselung von Thema und Abgabedatum einer Masterarbeit.

LITERATUR

Davenport, T.H. 2005. Thinking for a living: how to get better performances and results from knowledge workers. [Buch]. Harvard Business Press, Boston.

De Giacomo, G. and Dumas, M. and Maggi, F. M. and Montali, M. 2015. "Declarative process modeling in BPMN". In *Proceedings of International Conference on Advanced Information Systems Engineering*. Springer, Cham. 84-100.

Fahland, D. and Lübke, D. and Mendling, J. and Reijers, H. and Weber, B. and Weidlich, M. and Zugal, S. 2009. "Declarative versus imperative process modeling languages: The issue of understandability". In *Enterprise, Business-Process and Information Systems Modeling*. Springer, Berlin, 353-366.

Hinkelmann, K. and Pierfranceschi, A. 2014. "Combining process modelling and case modelling". In *Proceedings of 8th International Conference on Methodologies, Technologies and Tools enabling e-Government MeTTeG14*, 83.

Hollingsworth, D., and Hampshire, U. K. 1995. *Workflow management coalition: The workflow reference model. Document Number TC00-1003*, 19, No. 16, 224.

OMG. 2013. *Business Process Model and Notation (BPMN), Version 2.0.2*.

OMG. 2016. *Case Management Model and Notation (CMMN) Version 1.1*.

Pesic, M. 2008. *Constraint-based workflow management systems: shifting control to users*. Technische Universiteit Eindhoven.

Reijers, H. A. and Rigter, J. H. M. and van der Aalst, W. M. 2003. The case handling case. *International Journal of Cooperative Information Systems*, 12, No. 3, 365-391.

Ter Hofstede, A. H. and Van der Aalst, W. M. and Adams, M. and Russell, N. 2009. *Modern Business Process Automation: YAWL and its support environment*. Springer Science & Business Media.

van Der Aalst, W. M. and Pesic, M. and Schonenberg, H. 2009. "Declarative workflows: Balancing between flexibility and support". *Computer Science-Research and Development*, 23, No. 2, 99-113.

KONTAKT

Prof. Dr. Peter Edelmann
Technische Hochschule Mittelhessen
Wilhelm-Leuschner-Straße 13
61169 Friedberg
Peter.Edelman@mnd.thm.de

Prof. Dr. Armin Wagenknecht
Technische Hochschule Mittelhessen
Wiesenstraße 14
35396 Gießen
Armin.Wagenknecht@mni.thm.de

M.Sc. Michel Jonathan Schmitz
Technische Hochschule Mittelhessen
Wiesenstraße 14
35396 Gießen
michel.jonathan.schmitz@its.thm.de

CHANGE DETECTION FOR AREA SURVEILLANCE
USING A MOVING CAMERA

Tatsuhisa Watanabe, Tomoharu Nakashima, and Yoshifumi Kusunoki

Graduate School of Humanities and Sustainable System Sciences

Osaka Prefecture University

Gakuen-cho 1-1, Sakai, Osaka 599-8531, Japan

Email: {tatsuhisa.watanabe, tomoharu.nakashima, yoshifumi.kusunoki}@kis.osakafu-u.ac.jp

KEYWORDS

Change detection, Area surveillance, Monocular camera, Autonomous robot.

ABSTRACT

This paper tackles area surveillance with a moving camera by change detection. None of the existing datasets for change detection meets a surveillance scenario where a camera is mounted on a moving platform and pointed in the direction of moving. Thus, this paper creates a new dataset including several challenging points. For this dataset, this paper employs a composable method and proposes some components. To evaluate the proposed components, some corresponding classic methods were also tested on the dataset. As a result, the proposals outperformed them. Moreover, this paper investigated the relationship between the parameters of the components and their performance.

INTRODUCTION

Surveillance systems have been attracting attention because they have a great potential to reduce the workload of monitors. Surveillance systems can be applied to various fields such as urban monitoring, agriculture, and traffic analysis with manifold sensors. In recent years, some sensors have been mounted on Unmanned Aerial Vehicles or Unmanned Ground Vehicles (UGVs) as the development of industrial technologies. This paper aims to automatically observe areas for security using these autonomous vehicles. To do so, this paper overviews some previous methods and datasets in the following paragraphs and proposes a new dataset and method.

Regarding the area monitoring, there are two types of automatic methods: target-limited and target-agnostic. The target-limited methods focus on a specific anomaly including human behaviors (Morais et al. 2019; Singh et al. 2018). While the target-limited methods performed well, they cannot be a complete replacement for humans because they can only detect the expected target. The idea of combining them does not work because one cannot obtain or even list all possible abnormal patterns. The target-agnostic methods assume the available data as a distribution of normal situations and detect samples far from it as anomalies (Chu et al. 2019; Hao et al. 2019).

The target-agnostic methods for area surveillance can be roughly divided into two groups based on situation types. One group is for the place where people are NOT supposed to be. Methods of this sort have to be able to detect the emergence or disappearance of anything. The other group is for the place where people appear. In such places, detectors are required to report anomalous behaviors of humans too. The former situation can be solved by change detection (CD) and the latter by anomaly detection. The former is more important in practice because if there are people, they can take action.

Although a large number of studies devised CD methods, all of them employed fixed cameras. While research of this kind plays an important role in surveillance, blind spots of the fixed cameras can arouse a controversy over security. The blind spots can be reduced by mounting cameras on a moving platform such as autonomous vehicles.

There are only four datasets for CD with moving vehicles. One dataset, known as VDAO (Silva et al. 2014), was created inside an offshore facility. A camera on a mobile robotic platform on a straight rail was used. This dataset contains 15 different abnormal objects such as bags. (Sakurada and Okatani 2015) constructed two datasets consisting of panoramic images: TSUNAMI and GSV. TSUNAMI captured scenes of tsunami-damaged areas in Japan. GSV is a collection of images on Google Street View. The last dataset is the so-called VL-CMU-CD dataset (Alcantarilla et al. 2018). It includes pictures taken in the city of Pittsburgh, PA, USA, over a year.

The four CD datasets with moving devices do not suffice for area surveillance. Every dataset but VDAO (i.e. TSUNAMI, GSV, and VL-CMU-CD) was not designed for the field of surveillance. Thus, their change types such as buildings are not anomalous. The VDAO dataset only contains abandoned objects, not humans. Moreover, they do not contain looming motion. Such motion is unavoidable if one employs a moving camera in a narrow place including a hallway. In response to the inadequacy of the existing datasets, a new CD dataset is created with a moving monocular camera on a hallway. Compared to the existing datasets, the proposed one has some challenging points: 1) looming motion, under which everything varies grad-

ually in size and position; 2) non-identical trajectory and inconsistent viewing angles, which cause parallax leading to false alarm; 3) illumination change due mainly to different times of the day. The dataset is detailed more in the EXPERIMENTS AND RESULTS section. To tackle the proposed dataset, a composable procedure is employed as in (Carvalho et al. 2019). However, it was proposed for the VDAO dataset, so three new components are designed for the proposed dataset: Video Compression (VC), Temporal Alignment (TA), and Frame Comparison (FC).

To evaluate the proposed TA and FC components, classic TA and FC methods are also tested. (Evangelidis and Bauckhage 2013) proposed a TA method using local descriptors. One of the TA datasets they tackled is similar to the proposed dataset, which contains looming motion. (Carvalho et al. 2019) proposed a structured method for the VDAO dataset and employed Zero-mean Normalized Cross Correlation (ZNCC) for dissimilarity calculation.

The contributions of this paper are two-fold.
- A dataset has been created with a moving monocular camera. This is the first video-CD dataset that has looming-motion for area surveillance.
- A structured way has been proposed to deal with the proposed dataset.

METHOD

This paper employs a composable procedure as in (Carvalho et al. 2019) with three new components. Figure 1 illustrates its whole procedure, where dotted line boxes indicate that the corresponding part was originally proposed in the literature, and filled-in boxes are the new components that this paper proposes. The input is a pair of videos: reference video and target one. The method detects changes in the target video against the reference one. The first process, VC, is a novel process for reducing the computation cost of the downstream processes. TA synchronizes a compressed version of the target video to that of the reference one. SA performs image registration between each target frame and the matched reference one. FC computes dissimilarity values of each matched frame pair and binarizes them with a threshold.

For each component, visual information of videos needs to be extracted. To do so, famous CNN architectures, called VGG13 and VGG16 (Simonyan and Zisserman 2014), are used. They consist of three types of layers: convolution, max pooling, and fully connected layer. In this paper, all the fully connected layers of VGGs are discarded because the components demand just spatial information, not features for classification. Training a deep network requires a considerable amount of data and time. Therefore, this paper exploits the pre-trained parameters of VGGs on the ImageNet dataset (Deng et al. 2009). For VC and TA, the last pooling layer of VGG16 is replaced with global average pooling (GAP) (Lin et al. 2013). GAP is considered to lose spatial information. However, it seems that this

weakness potentially renders a matching method insensitive to parallax or too partial textures. The GAP version of VGG16, noted as VGG16', returns a C dimensional feature map for the input image.
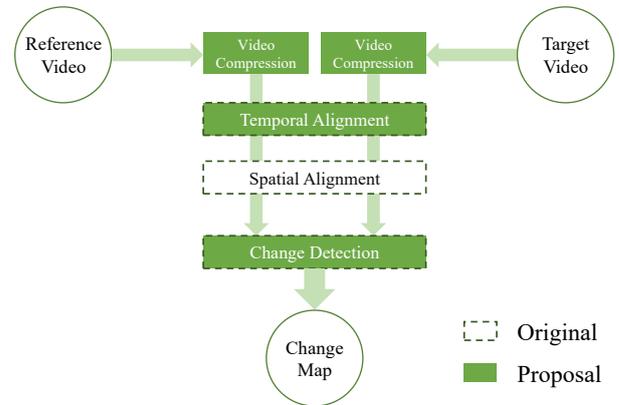
Figures 1: The Whole Procedure of the Proposed Method

Proposed Video Compression (VC)

With the aim of area surveillance, it is not necessary to inspect all frames since several consecutive frames contain almost the same contents. Therefore, VC reduces redundant frames according to the similarity of sequential ones. Let $S^r = \{I_1^r, I_2^r, ..., I_{n_r}^r\}$, $S^t = \{I_1^t, I_2^t, ..., I_{n_t}^t\}$ be the reference sequence and the target one, respectively. n_r and n_t are the number of the reference frames and that of the target ones, respectively. The aim of this component is to retrieve key frame indices $X^r = \{i_1^r, i_2^r, ..., i_{n_{rc}}^r\}$, $X^t = \{i_1^t, i_2^t, ..., i_{n_{tc}}^t\}$. n_{rc} and n_{tc} are the number of the reference key frames and that of the target ones, respectively. Unless otherwise noted, all the frames but the key frames are to be disregarded in the ensuing processes. VC is performed in a chronological manner as in Algorithm 1, where $cos\_sim(v_1, v_2)$ is the cosine similarity value of two vectors v_1, v_2. Note that this step processes independently the target video and the reference one.

Algorithm 1 Video Compression

In: S, T_c // S: set of video frames. T_c: threshold.
Out: X // X: set of key frame indices.
1:    array $X \leftarrow \{1\}$
2:    $l \leftarrow 1$
3:    **while** $l < S.length$ **do**
4:      **for** $h = l + 1$ to $S.length$ **do**
5:        **if** $cos\_sim(VGG16'(S[l]), VGG16'(S[h])) < T_c$ **then**
6:          X.append(h)
7:          break
8:        **end if**
9:      **end for**
10:     $l \leftarrow h$
11: **end while**
12: **return** X

Proposed Temporal Alignment (TA)

TA matches frames in the target video to those in the reference video. This paper proposes a new TA method using deep features. Let $\hat{X} = \{\hat{k}_1, \hat{k}_2, ..., \hat{k}_{n_{tc}}\}$ be the indices of the matched reference frames. $\hat{X}$ is calculated by Equation (1).

$$\hat{k}_l = \arg\max_h \; cos\_sim(VGG16'(I_{i_l^t}^t), VGG16'(I_{i_h^r}^r)),$$
$$(1)$$

where $l = 1, 2, ..., n_{tc}$. To prevent abruption from the previously matched index, a search range is restricted as $h \in \{m_l^{back}, m_l^{back} + 1, ..., m_l^{front}\}$. m_l^{back}, m_l^{front} are calculated by Equations (2), (3), respectively.

$$m_l^{back} = \begin{cases} i_1^r & \text{if } l = 1, \\ \max(i_1^r, \hat{k}_{l-1} - back) & \text{otherwise,} \end{cases} \quad (2)$$

$$m_l^{front} = \begin{cases} front & \text{if } l = 1, \\ \min(i_{n_{rc}}^r, \hat{k}_{l-1} + front) & \text{otherwise,} \end{cases} \quad (3)$$

where $back$ and $front$ are hyperparameters, discussed in the EXPERIMENTS AND RESULTS section.

Spatial Alignment (SA)

It is impossible to take all videos in an exactly consistent angle or position. Consequently, even if TA perfectly synchronizes the input videos, the matched frame pairs will have different image planes. Thus, SA, or image registration, is performed with Homography transformation. Let $H(I_1, I_2)$ be a transformed image of I_1 to I_2. The outside of an image plane is treated as black (i.e. RGB value (0, 0, 0)), when it gets into the image plane by Homography transformation. This black area would be detected as change in the following stage. Therefore, the counterpart of the target frame is masked. I' signifies a masked version of an image I.

Proposed Frame Comparison (FC)

The last component of the proposed method is FC. FC compares the spatiotemporally aligned frame pairs in the upstream processes and calculates dissimilarity maps. After that, it computes change maps by binarizing the dissimilarity maps with a threshold. Dissimilarity maps are obtained based on the idea by (Kim et al. 2017). They proposed a similarity calculation method with the VGG13 network for template matching. Feature maps of a template image and a target one were extracted from a mid-convolutional-layer of the network. The target feature map was searched with a sliding window in order to determine the patch of the target image most similar to the template image. They used Normalized Cross Correlation (NCC) as a similarity criterion.

A feature map of the m-th target frame and that of the matched reference frame are denoted as $M_m^t = VGG13((I_m^t)')$, $M_m^r = VGG13(H(I_{\hat{k}_m}^r, I_m^t))$, respectively. Note that the target frames and the reference ones,

in the setting of this paper, have the same resolution, meaning their feature maps are of the same shape. This fact enables the feature maps to be compared in a position-wise manner as $SM_{m,i,j} = cos\_sim(M_{m,i,j}^t, M_{m,i,j}^r)$, where $i \in \{1, 2, ..., H\}$, $j \in \{1, 2, ..., W\}$, and $M_{m,i,j}^t$ and $M_{m,i,j}^r$ are C-dimensional vectors. It is noteworthy that NCC corresponds to cosine similarity when the target pair is two vectors. By following (Kim et al. 2017), one can obtain a similarity map since it was proposed for template matching. Thus, it is converted into a dissimilarity map as $DM_{m,i,j} = 1 - SM_{m,i,j}$.

A multi-scale option is introduced as in (Carvalho et al. 2019) with some modifications. (Carvalho et al. 2019) obtained different-scale maps by resizing an frame. Subsequently, they resized them to the input size and just added them up. This way can be followed, but there are some constraints because of the CNN attribution. Some CNN layers downsample an image. While their processing, they would discard the right-end or bottom-end information of the input image not even considering it due to the filter size or the stride of those layers. VGG13 has five pooling layers, the window size of which is 2×2. The other layers of VGG13 do not affect the output size. For this reason, the resolution of the input should be divisible by $2^5 = 32$ to avoid loss of spatial information. Moreover, the aspect ratio of the proposed dataset is 16:9. Putting these conditions together, two resolution candidates are obtained: 512×288 and 1024×576. This paper extracts features only from the final pooling layer of VGG13. The feature maps from it contain the most abundant peripheral context than those from the preceding layers.

Three weight types are proposed to combine different-scale dissimilarity maps DM^k, $k \in \{1, 2, ..., n_{dm}\}$. n_{dm} denotes the total number of DM. Equation (4) shows how to create a weighted map, $weighted\_DM$.

$$weighted\_DM_{i,j} = \sum_k w^k r DM_{i,j}^k, \quad (4)$$

where rDM^k is the resized DM^k to the input size (W^{org}, H^{org}) with nearest neighbor interpolation. $i \in \{1, 2, ..., W^{org}\}$ and $j \in \{1, 2, ..., H^{org}\}$ are xy-coordinate positions. w^k is the k-th weight. One weight type is MAX as in Equation (5).

$$w^k = \frac{\max(DM^k)}{\sum_{l=1}^{n_{dm}} \max(DM^l)}. \quad (5)$$

A change is more detectable by a suited-scale map than the other scale maps. Thus, using a certain-scale map probably results in higher dissimilarity values for the corresponding-scale change than the other scale maps. Based on this idea, the MAX weight type is designed not to miss changes. Another is EQUAL as in Equation (6).

$$w^k = \frac{1}{n_{dm}}. \quad (6)$$

In EQUAL, all weights have the same value. The third weight type is LARGE as in Equation 7. Assume the

size of DM^1 be the smallest of the maps DM_1, DM_2, ..., $DM_{n_{dm}}$ and set w^1 as the possible maximum weight.

$$w^k = \begin{cases} \dfrac{1}{1 + \sum_{l=2}^{n_{dm}} \max(DM^l)} & \text{if } k = 1, \\ \dfrac{\max(DM^k)}{1 + \sum_{l=2}^{n_{dm}} \max(DM^l)} & \text{otherwise.} \end{cases} \quad (7)$$

With the LARGE weight, the smallest dissimilarity map is assigned with the possible maximum weight. In other words, the weights of the other scale maps would have smaller weights than the case of the other weight types. The smallest map contains the spatially roughest information, meaning it includes less environmental effects such as parallax than the other maps. Thus, the LARGE weight is expected to mitigate environmental effects causing false alarm. Once the weighted map is obtained by Equation (4), a change map can be calculated by binarizing the weight map. The threshold value is discussed in the EXPERIMENTS AND RESULTS section.

EXPERIMENTS AND RESULTS

Dataset

The existing datasets do not contain anomalous changes. Therefore, a new dataset has been created by recording some looming-motion videos with a radio-controlled vehicle. The vehicle ran on a straight corridor at a speed of 0.5 meters per second, and never moved backward. Its trajectories were not identical, and the viewing angle of the vehicle was inconsistent. The proposed dataset includes two sets of a reference and six target videos about 1.5 minutes long each, so the total number of the videos is fourteen. The two sets were captured at different times of the day: day and night. Table 1 shows what kind of changes the target videos contain. Each of the target videos was temporally, spatially aligned to the time-wise corresponding reference video. For TA assessment, each target frame was temporally aligned to a reference one at hand. Subsequently, dissimilarity maps were calculated by comparing each of the aligned frame pairs. To evaluate FC performance, each change was labeled with a bounding box. The proposed dataset is challenging due mainly to parallax or strong illumination change.

Parameter setting

The proposed method has some adjustable parameters. T_c was set to 0.995. For TA, $back$ was fixed to zero since the camera never moved backward in the dataset. Preferable values for $front$ were roughly searched for by grid search with a set of values (3, 5, 7, 10). Consequently, this paper chose $front= 7$ for the day targets and $front= 3$ for the night ones. As referred to in the METHOD section, a multi-scale option was employed, and the input images were resized to two scales: 512×288 and 1024×576. For ZNCC, this paper followed (Carvalho et al. 2019) and prepared scales: 20×11, 40×22, 80×45, and 160×90. Besides, another scale 320×180 was also tested for a deeper survey. The window size of ZNCC was set to five.

Table 1: Change Types in the Proposed Dataset

time	data name	included change type
day	fallen	person (fallen)
		bottle
	standing	person (standing)
		umbrella (dropped)
	walking	person (walking)
		umbrella (leaning)
		door
	stacked	two boxes (stacked)
	separate	two boxes (separate)
	bag	bag
night	fallen	person (fallen)
		bottle
		shoe
	standing	person (standing)
		umbrella (leaning)
		shoe
	walking	person (walking)
		umbrella (dropped)
		shoe
	stacked	two boxes (stacked)
	separate	two boxes (separate)
	bag	bag

Experiment

The proposed TA method was compared with (Evangelidis and Bauckhage 2013). To give a quantitative comparison, this paper followed (Diego et al. 2013). They set a ground-truth interval $[l_t, u_t]$ for each index t of a sequence and calculated TA errors as in Equation (8).

$$err(t, \hat{k}_t) = \begin{cases} 0 & \text{if } l_t \leq \hat{k}_t \leq u_t, \\ \min(|l_t - \hat{k}_t|, |u_t - \hat{k}_t|) & \text{otherwise,} \end{cases} \quad (8)$$

where $\hat{k}_t$ is the t-th index matched by a TA method and $t \in \{1, 2, ..., n_{rc}\}$. The one-by-one ground truth GT_t, which the proposed dataset included, was expanded by one on the negative and positive sides (i.e. $l_t = \max(1, GT_t - 1)$, $u_t = \min(n_{rc}, GT_t + 1)$). Table 2 shows rates of frames with equal or less than each error. TA with VGG16' provided better results than (Evangelidis and Bauckhage 2013) in all the videos. In case of $err = 0$, the error gaps are at least 8.9% (day/separate) and at most 54.5% (day/fallen). Even when comparing the $err = 0$ results by VGG16' and the $err \leq 2$ results by (Evangelidis and Bauckhage 2013), most of the former results are better. Comparing the day with the night, both methods rather struggled to align the day sequences. Looking at $err = 0$, the gaps between day/Average and night/Average are 14.9% (VGG16') and 39.6% (Evangelidis and Bauckhage 2013). This is because sunlight through windows formed different shapes on the wall and floor and affected the surrounding brightness, making the day videos of the proposed dataset more challenging. This sunlight worsened (Evangelidis and Bauckhage 2013) more strongly than VGG16' because the former method using local de-

scriptors unfortunately captured the sunlight changing its form. On the other hand, such local changes were invisible for VGG16' thanks to its GAP.

With the TA results by VGG16', the FC performance of the proposed pipeline was evaluated by the area under the receiver-operator curve (AUC). VGG13 and ZNCC with the three weight types were compared as shown in Table 3. This result only exhibits the best combination of scales: [512×288, 1024×576] for VGG and [20×11, 40×22, 80×45] for ZNCC. Table 3 indicates three notable points. Firstly, VGG13 outperformed ZNCC in all of the scenarios. Secondly, both methods performed the worst on day/walking and night/stacked for each time. What deteriorated them is investigated in the following paragraph. Thirdly, the weight type provided just a marginal difference. This is because if just a single pixel in one of the different-scale maps has a high value, it pushes up the weight of that map. Therefore, all weights ended up getting almost the same value.

Table 4 shows AUC scores of the FC methods with and without the proposed TA results. Only the LARGE weight type was shown as it performed the best. At day/walking and night/stacked in Table 4, large gaps can be seen. This suggests the failure of TA led to the terrible FC performance. This suggestion was confirmed by counting detectable pixels, which were non-masked ones in the METHOD section. The inner rate columns in Table 4 show each detectable pixel rate (%). Non-change areas are the outside of bounding boxes, and change areas are the inside. As one can see, the change inner rates for day/walking and night/stacked are obviously low, meaning a large part of the change areas was regarded as unchanging. Therefore, Table 4 proves TA plays a pivotal role in CD.

This paper expanded on how scale sizes affected results. Table 5 shows AUC scores for VGG13 and ZNCC with single scales. Smaller resolutions provided better scores because the AUC was a pixel-wise criterion. That is, a method tuned for larger objects contributes to the score more than smaller ones. Also, this tendency can be found in the best set of scales for ZNCC ([20×11, 40×22, 80×45]). Another notable point is that, comparing the AUC scores in Table 5 with the AUC scores using the TA ground truth in Table 4, the combination of multi scales enhanced the detection ability. In addition to the AUC, this paper looked into the relationship of scales and dissimilarity values for each object. This paper calculated the median of dissimilarity values belonging to each change type or the background and then a ratio of each object median to the background one. If a ratio is less than 1.0, the corresponding change is indistinguishable from the background. The higher it is, the more detectable the change is. Note that the looming motion in the videos significantly varies the size of changes. Thus, Table 6 only shows "person" and "bottle", a large and a relatively small change, in {day, night}/fallen. One can see the tendency of larger

resolutions spotting smaller changes and vice versa.

Finally, to discuss the FC performance for each change type, a ratio of each object median to the background one was computed with VGG13 ([512×288, 1024×576]) and ZNCC ([20×11, 40×22, 80×45]) as shown in Table 7. The weight type was fixed to LARGE as in Table 4. Table 7 indicates some characteristics of ZNCC and VGG13. ZNCC shows distinctively strong and weak points. It failed to detect the smallest change, "bottle". Moreover, the value for a relatively small object "umbrella" is significantly smaller than the other changes except for "door". Note that although "shoe" might sound small, it appears close to the vehicle trajectories. Thus, "shoe" looks big in the proposed dataset. On the other hand, VGG13 successfully detected "bottle". Its ratio is actually close to 1.0, but this result seems reasonable because "bottle" is not only small but also unobtrusive in the proposed dataset. For the other changes including "umbrella", VGG13 almost impartially spotted them. This implies that VGG13 does not largely depend on the input scales. This is because its convolutional layers acquire surrounding information.

CONCLUSION

This paper aims to automatically monitor areas for security using a moving camera instead of humans. None of the existing CD datasets was designed for such a purpose. Thus a new dataset for area surveillance has been built with a UGV. Subsequently, this paper has introduced a structured method and devised three components for it: VC, TA, and FC. For FC, three ways to combine different-scale maps have also been proposed. To perform an evaluation, the proposed TA and CD methods were compared with classic methods. Through the experiments, this paper showed the effectiveness of the proposed method in area surveillance using a moving camera.

There are some limitations in the proposed method. First, the proposed CD method cannot detect changes in a target frame if the matched reference frame does not contain the spatially corresponding region. In terms of false positive, there were some times the method falsely detected objects as changes due to difference in viewing angle or position and the sunlight. Second, the FC performance strongly depends on the preceding procedure: TA and SA, as shown in Table 4. Third, if changes appear in a dominant part of an image, TA and SA would provide a poor result. Finally, the reference video has to contain the whole scenes of the target video. This limits a range of applications.

A piece of the future work is to improve the proposed method by overcoming the limitation. It is necessary to research how to make methods robust to environments. Evaluation-wise, this paper performed an evaluation with the pixel-wise AUC. As aforementioned, it tended to give better scores to a method tuned for larger changes. This tendency is not appropriate for surveillance. For this reason, a new frame-level evaluation should be consid-

Table 2: Frame Rates (%) with Equal or Less than Each Error for TA

time	data	VGG16'			georgios		
		$err = 0$	$err \leq 1$	$err \leq 2$	$err = 0$	$err \leq 1$	$err \leq 2$
day	fallen	68.1	80.1	84.3	13.6	20.4	26.7
	standing	52.3	70.5	76.7	8.0	11.9	14.8
	walking	55.7	73.4	83.7	18.2	26.1	34.5
	stacked	78.7	88.1	94.1	38.6	46.0	49.0
	separate	78.8	90.2	90.7	69.9	85.0	89.1
	bag	67.7	80.8	83.8	18.2	23.7	27.8
	Average	66.9	80.5	85.6	27.8	35.5	40.3
night	fallen	91.3	96.9	100.0	81.6	87.2	91.3
	standing	95.3	97.9	98.4	68.6	79.1	84.3
	walking	98.5	100.0	100.0	85.8	92.9	93.9
	stacked	79.7	91.9	97.7	54.7	65.7	70.9
	separate	68.0	74.6	80.1	27.6	40.3	51.9
	bag	97.8	100.0	100.0	86.0	91.6	98.3
	Average	81.8	93.6	96.0	67.4	76.1	81.8

Table 3: AUC Scores with the Proposed TA Component for FC

time	data	VGG13			ZNCC		
		MAX	EQUAL	LARGE	MAX	EQUAL	LARGE
day	fallen	0.870	**0.871**	**0.871**	0.745	0.739	0.738
	standing	**0.842**	**0.842**	**0.842**	0.731	0.730	0.729
	walking	0.686	0.688	**0.689**	0.654	0.661	0.663
	stacked	0.860	0.862	**0.863**	0.708	0.709	0.704
	separate	0.813	0.817	**0.818**	0.765	0.784	0.786
	bag	**0.926**	0.925	0.925	0.812	0.803	0.801
	overall	0.833	0.834	**0.835**	0.736	0.738	0.737
night	fallen	0.913	0.914	**0.914**	0.834	0.837	0.838
	standing	0.915	**0.916**	**0.916**	0.821	0.820	0.820
	walking	0.917	0.921	**0.922**	0.864	0.877	0.879
	stacked	**0.584**	0.583	0.583	0.512	0.513	0.515
	separate	**0.899**	**0.899**	**0.899**	0.837	0.840	0.841
	bag	0.901	0.903	**0.904**	0.808	0.812	0.811
	overall	0.855	**0.856**	**0.856**	0.779	0.783	0.784
Average		0.844	**0.845**	**0.845**	0.758	0.760	0.760

Table 4: Change/Non-Change Inner Rates (%) and AUC Scores with and without the Proposed TA Component

time	data	TA_VGG16'				TA_GROUND_TRUTH			
		non-change inner rate	change inner rate	VGG13	ZNCC	non-change inner rate	change inner rate	VGG13	ZNCC
day	fallen	83.2	97.1	0.871	0.738	86.9	98.6	**0.908**	0.770
	standing	80.2	92.9	0.842	0.729	85.4	90.8	**0.869**	0.754
	walking	79.4	69.8	0.689	0.663	82.0	85.9	**0.831**	0.738
	stacked	86.7	100.0	**0.863**	0.704	88.3	100.0	0.853	0.691
	separate	92.3	86.2	**0.818**	0.786	93.6	81.1	0.798	0.786
	bag	81.4	97.9	0.925	0.801	83.5	100.0	**0.966**	0.862
night	fallen	92.6	98.5	**0.914**	0.838	92.5	97.9	**0.914**	0.842
	standing	87.9	99.6	0.916	0.820	88.1	98.1	**0.929**	0.841
	walking	91.2	95.1	0.922	0.879	90.5	97.3	**0.936**	0.889
	stacked	86.7	64.9	0.583	0.515	88.5	98.1	**0.861**	0.727
	separate	82.5	98.6	0.899	0.841	87.3	99.4	**0.940**	0.869
	bag	93.5	93.4	0.904	0.811	94.2	93.4	**0.911**	0.815
Average				0.845	0.760			**0.893**	0.799

Table 5: AUC Scores for Different Scales with the TA Ground Truth

time	data	VGG13		ZNCC				
		512×288	1024×576	20×11	40×22	80×45	160×90	320×180
day	fallen	0.910	0.880	0.703	0.766	0.764	0.738	0.699
	standing	0.865	0.841	0.706	0.741	0.719	0.683	0.631
	walking	0.848	0.790	0.731	0.725	0.678	0.618	0.571
	stacked	0.837	0.831	0.572	0.689	0.717	0.719	0.697
	separate	0.815	0.757	0.834	0.758	0.677	0.607	0.550
	bag	0.952	0.949	0.762	0.842	0.863	0.824	0.754
	overall	0.871	0.841	0.718	0.753	0.736	0.698	0.650
night	fallen	0.908	0.890	0.807	0.830	0.818	0.768	0.710
	standing	0.919	0.905	0.787	0.835	0.804	0.704	0.609
	walking	0.936	0.895	0.891	0.865	0.819	0.740	0.663
	stacked	0.833	0.854	0.686	0.709	0.726	0.728	0.699
	separate	0.936	0.918	0.847	0.856	0.834	0.796	0.747
	bag	0.909	0.880	0.793	0.815	0.793	0.707	0.599
	overall	0.907	0.890	0.802	0.818	0.799	0.740	0.671
Average		0.889	0.866	0.760	0.786	0.768	0.719	0.661

Table 6: The Ratio of Each Median of Two Objects to the Background One for Each Scale

time	change	VGG13		ZNCC				
		512×288	1024×576	20×11	40×22	80×45	160×90	320×180
day	person	3.39	2.76	1.90	4.36	4.80	3.58	2.54
	bottle	2.78	2.64	1.20	1.09	3.40	2.54	2.14
night	person	4.95	3.13	13.50	20.00	13.67	5.41	2.64
	bottle	0.92	1.64	0.50	0.33	0.33	1.29	1.15

Table 7: The Ratio of Each Object Median to the Background One

	person	bottle	umbrella	door	box	bag	shoe
VGG13	3.23	1.46	2.82	2.80	3.30	3.38	2.91
ZNCC	6.10	0.70	2.70	3.20	6.40	5.50	4.40

ered. Finally, the proposed dataset contains little variation. Thus, it is required to record videos in different seasons or weather. On top of that, other places such as curves should be included.

REFERENCES

Alcantarilla, P. F.; S. Stent; G. Ros; R. Arroyo; and R. Gherardi. 2018. "Street-view change detection with deconvolutional networks", *Journal of Autonomous Robots*, Vol. 42 (May), 1301-1322.

Carvalho, G. H. F. de; L. A. Thomaz; A. F. da Silva; E. A. B. da Silva; and S. L. Netto. 2019. "Anomaly Detection with a Moving Camera using Multiscale Video Analysis", *Journal of Multidimensional Systems and Signal Processing*, Vol. 30, Issue 1 (January), 311-342.

Chu, Wenqing; H. Xue; C. Yao; and D. Cai. 2019. "Sparse Coding Guided Spatiotemporal Feature Learning for Abnormal Event Detection in Large Videos", *IEEE Transactions on Multimedia*, Vol. 21, Issue 1, 246-255.

Deng, J.; W. Dong; R. Socher; L.-J. Li; K. Li; and L. Fei-Fei. 2009. "ImageNet: A large-scale hierarchical image database", *Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition*, 20-25.

Diego, F.; J. Serrat; and A. M. López. 2013. "Joint Spatio-Temporal Alignment of Sequences", *Journal of IEEE Transactions on Multimedia*, Vol. 15, No. 6 (October), 1377-1387.

Evangelidis, G. D. and C. Bauckhage. 2013. "Efficient Subframe Video Alignment using Short Descriptors", *Journal of IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 35, No. 10 (October), 2371-2386.

Hao, Y.; Z.-J. Xu; Y. Liu; J. Wang; and J.-L. Fan. 2019. "Effective Crowd Anomaly Detection through Spatio-temporal Texture Analysis", *Journal of Automation and Computing*, Vol. 16, Issue 1 (February), 27-39.

Kim, J.; J. Kim; S. Choi; M. A. Hasa; and C. Kim. 2017. "Robust Template Matching using Scale-Adaptive Deep Convolutional Features", *Proceedings of Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, 708-711.

Lin, M.; Q. Chen; and S. Yan. 2014. "Network In Network", *Proceedings of International Conference on Learning Representations*, 10 pages.

Morais R.; V. Le; T. Tran; B. Saha; M. Mansour; and S. Venkatesh. 2019. "Learning Regularity in Skeleton Trajectories for Anomaly Detection in Videos", *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*), 11996-12004.

Sakurada, K. and T. Okatani. 2015. "Change detection from a street image pair using CNN features and superpixel segmentation", *Proceedings of British Machine Vision Conference*, 12 pages.

Silva, A. F. da; L. A. Thomaz; G. Carvalho; M. T. Nakahata; E. Jardim; J. F. L. de Oliveira; E. A. B. da Silva; S. L. Netto; G. Freitas; and R. R. Costa. 2014. "An Annotated Video Database for Abandoned-Object Detection in a Cluttered Environment", *Proceedings of 2014 International Telecommunications Synposium*, 5 pages.

Simonyan, K. and A. Zisserman. 2014. "Very deep convolutional networks for large-scale image recognition", *Proceedings of International Conference on Learning Representations*, 14 pages.

Singh, A.; D. Patil; and S. N. Omkar. 2018. "Eye in the Sky: Real-Time Drone Surveillance System (DSS) for Violent Individuals Identification Using ScatterNet Hybrid Deep Learning Network", *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (*CVPRW*), 1710-1718.

Implementation of Robotic Process Automation
in a Retail Company

Franziska Kaiser

Pforzheim University
Tiefenbronner Straße 65
75175 Pforzheim
kaiserff@hs-pforzheim.de

Peter Trkman

University of Ljubljana
Kardeljeva ploščad 17
1000 Ljubljana
peter.trkman@ef.uni-lj.si

Frank Morelli

Pforzheim University
Tiefenbronner Straße 65
75175 Pforzheim
frank.morelli@hs-pforzheim.de

KEYWORDS

Business Process Management, Robotic Process Automation, Critical Success Factors

ABSTRACT

In recent years, Robotic Process Automation (RPA) has emerged as a tool to automate processes to bring higher productivity, process speed, lower costs, and errors in process execution. However, many companies are uncertain on if, how, and when to use RPA. Therefore, the purpose of this paper is to explain what RPA is and to provide core messages for companies on:
1. How to choose suitable processes for RPA,
2. What to look out for during implementation, and
3. Which factors can influence the successful introduction of the technology.
The conduction of a case study in a German retail company helped to reveal this information.

INTRODUCTION

Every company has business processes that need to be done regularly. Depending on the process, the activities within a process can be complex or relatively simple and repetitive. Especially for those simple, structured, non-critical processes, human intervention is not necessarily needed (Power et al., 2017). Extracting data from an Enterprise Resource Planning (ERP) system, export it to Excel, and process the data to create a weekly or daily report is a typical example. Since continuous improvement is crucial for companies, it is a major challenge to always think of ways on how to improve the business processes (Vanwersch et al., 2016). New technologies emerging due to digitalization offer support to exploit given improvement potentials.

One improvement idea is the introduction of RPA. RPA is an umbrella term for software tools to automate repetitive and simple processes by programming bots to act in a way just as humans would do (Gartner, n.d.). RPA's goal is to picture tasks initially done by humans 1:1 by bots. Instead of employees manually writing data entries in, e.g., ERP systems, the bots are programmed to do so. Through the usage of surface automation, RPA can be implemented quickly without major changes to any application systems used by the company. Furthermore,

RPA may lead to fewer mistakes made by humans, higher process throughput, and consequently increased efficiency (Santos et al., 2020).

RPA bots can either run attended or non-attended. Attended bots run on the desktop of an employee. Usually, the employee develops the bot himself, triggers it when needed, and watches the bot execute the process steps. Non-attended bots run on virtual machines. The orchestration is organized centrally by using a cockpit (Langmann and Turi, 2020). The cockpit triggers the process execution. One main benefit of non-attended bots is that employees can fulfil other tasks while the bots execute the processes (Koch and Fedtke, 2020).

The importance of RPA is illustrated by the Gartner report that current spending by companies on RPA software was above $1.5 billion in 2020 and is predicted to expand to $2.9 billion in 2021 (Rauch, 2020). Further, Deloitte estimated that RPA will hit almost universal adoption in companies by 2023 (Casey, 2020). The COVID-19 pandemic caused the RPA demand to spike since companies were forced to work with less staff, cut employee hours and provide home office possibilities (Rauch, 2020).

Even though process automation is often perceived as a game-changer, RPA also has some drawbacks. One problem is that, in practice, RPA is used to automate inefficient processes by "patching" non-existent Application Programming Interfaces (APIs) through the usage of surface automation (Koch and Fedtke, 2020). If the risks of automation for the targeted processes are not evaluated in advance, the impact on the company can be tremendous (Power et al., 2017). Therefore, it is important to understand how to choose suiting processes for RPA (Santos et al., 2020).

Given the relative newness of the field, the factors that can influence the success of RPA projects are not well identified yet. The main research question of this paper is "Which factors should be considered when aiming to introduce RPA in a company or a department of a company?".

A case study conducted in a German retail company provides an analysis of the possibilities and outcome of RPA

by implementing non-attended automation for two defined logistics processes. Conducting the case study in the selected company, led to some lessons learned. These lessons learned may help other companies trying to implement RPA too. Therefore, these lessons learned are summarized in this paper as well as the benefits and drawbacks of an introduction to RPA. Finally, possible future developments of the technology are presented.

CASE STUDY

Methodology

The research strategy of this work is a case study. A case study is a deep analysis of a research object and its environmental conditions (Oehlrich, 2019). The research objects for this work are the logistics processes of the selected retail company.

For the discovery of the relevant processes in the logistics department, semi-structured interviews with one to two responsible employees of the department for each process were conducted. Depending on the complexity of the process, additional questions were asked. The goal of the interviews was to generate a detailed understanding of the process itself including the relevant systems which are used, the activities, and the exceptions within the process. Furthermore, the goal of the interviews was to enable the evaluation of whether RPA is a suitable automation method for the given process or not. Lastly, the manual execution of the process was recorded in a video during the interview. The recording serves as process documentation in case the bot crashes in the future and to help while developing the automation.

Existing process documentation within the department was used for process automation. Both processes were modelled with BPMN 2.0 notation which is generally understood by business users. BPMN 2.0 notation allows the creation and visualization of end-to-end processes by providing a set of rules and conventions for the model (vom Brocke and Rosemann, 2010).

AS-IS Situation

Firstly, a suitable RPA software provider (in this case UiPath) needed to be found. UiPath is the current market leader in RPA software solutions (Gartner, 2020). They gave a strong partner ecosystem with more than 250 technology partners, a strong development community, and many learning resources like online training (Ray et al., 2020). UiPath can also be used for other purposes like e.g., process mining or testing automation. Also, the RPA infrastructure, including the setup of virtual machines, purchase of licenses, etc. and the assignment of responsible employees for the RPA infrastructure, was done.

For attended bots, each employee in the company can download the UiPath Studio version and start programming. For non-attended bots, the RPA infrastructure needs to be used. The RPA infrastructure is managed through the UiPath Orchestrator. Employees, bots, and virtual machines can be assigned to department folders which are managed by the administrators. The employees within the folder can upload UiPath process files and start the processes. There is no company-wide concept for how the departments should organize process automation with RPA.

Project Execution

The BPM lifecycle is often used to structure and manage business processes (Bergener, et. al., 2019). It also improves the understanding of the role of technology within BPM (Dumas et al., 2013). Since RPA is a new technology, the BPM lifecycle provides a suitable structure for the different phases of the case study's project execution. The BPM lifecycle consists out of the following phases: process identification, discovery, analysis, redesign, implementation, monitoring, and control (Flechsig et al., 2019).
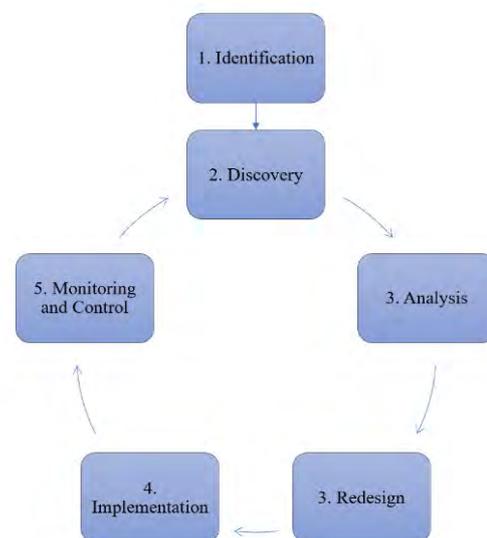
Identification
The first phase in the BPM lifecycle is the identification phase. However, the technology of RPA needs to be understood first. An initial kick-off meeting with the logistics employees introduced the idea of RPA in the department. Within the meeting, the definition of RPA has been introduced, as well as important characteristics for processes that can be automated with RPA, and the benefits of introducing RPA. Next, it was necessary to find suiting processes for RPA.

For the actual identification of RPA candidates, eight questions that should be answered with "yes" and two questions requesting additional information about a process have been defined. The questions aim to cover all must-criteria for RPA. The goal of the questionnaire was to identify two use cases where RPA can be a good solution. The questions are the following:

- Is the process rule-based? (It is possible to create a process model without too many exceptions.)
- Does the process include structured data only? (The input data is electronic and has a defined format, e.g., item number, prices, etc.?)
- Is the process repetitive? (The process is executed regularly, e.g., weekly or daily.)
- Is the process stable? (The process is not new and will not change soon.)
- Are no activities included which need to be done manually? (The process does not require any non-digital activities, e.g., signatures.)
- Is the process itself non-critical for the company? (If the bot crashes, this has no immediate impact.)
- Is it possible to provide a timeframe when the bot can run the process without interruptions?
- Does the process contain system breaks? (Are different tools and systems like, e.g., SAP, Excel and Email used in the process?)

Additional information:
- How complex is the process? (Low, middle, high)
- How long does the process execution need? (In hours per week)

The logistics department was asked to provide two processes for which the answer for each of the eight questions is yes, and the complexity of the process is rather low or middle.

Discovery
The next phase in the BPM lifecycle is the discovery phase. This phase thematizes understanding the processes in detail and create the AS-IS models. Therefore, one meeting per process has been scheduled to conduct semi-structured interviews. In the meetings, the manual execution of the processes has been recorded in a video. Also, any existing process documentation has been gathered, all questions about the processes have been clarified and eventual exceptions in the process execution have been documented.

The first process is called "Gross Load Preview". Figure 2 shows the BPMN model of the process. Since it contains three activities only, the process complexity is rather low. The entire process is executed in the logistics department.
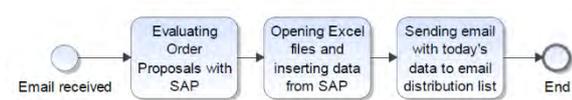
The process is executed daily by any employee from logistics, so it meets the requirement of repetitiveness. The employee needs about 10 minutes for the manual execution of the process. The process uses SAP, Excel and sends emails, so there are system breaks. The process is

stable, and it is completely rule-based. No signatures or similar activities are needed within the process. The process is not complex. The goal within the process is to evaluate the order proposal units for a picking date for each distribution center of the company. After evaluating the proposal units by using SAP, the numbers are sent to an email distribution list.

Each morning, at around 7:30 AM, the logistics department receives an automated email informing the employees about the completeness of the order data. This automated email is the trigger for the process to start. Next, an employee logs into the SAP system uses a transaction and selects a variant that prefills relevant data into the SAP input mask like the identification number for all relevant distribution centers. One difficulty within the input screen is to select the suiting picking date. Usually, the picking date should be the current day plus three workdays. However, the picking date can vary because one distribution center is in a region which is, e.g., on holidays. The employees within the logistics department are aware of these holidays per region and change the picking date manually. For the bot implementation, clear rules are needed. A solution was suggested to use another transaction that provides the delivery schedule for each distribution center. Within this transaction, it is possible to get the right picking date for each distribution center, copy it, and paste it in the order proposal transaction. After executing, the evaluation needs to be exported in an Excel file. Usually, the process needs to be executed from Monday to Friday. On the weekends, no data would be found in the relevant SAP transactions. However, it may happen that the bot needs to run on a Saturday. The company has some "logistics Saturdays" especially in November and December before Christmas. To integrate the "logictics Saturdays", the bot has been scheduled to run from Monday to Saturday. If the bot does not find any data, it is not supposed to crash but writes an email that it did not find any data due to holidays or no "logistics Saturday".

The next step is to open two Excel files that are used by the entire department. Next, the numbers from the SAP-export need to be copied into the corresponding fields in the Excel file. The corresponding field means that for each row in the SAP-export the number needs to be copied and pasted into the two Excel files when the source of supply and shipping point match with an entry in the Excel.

The last step in the process is to send an email to a defined list with an excerpt of the two Excel files showing the picking units for the most recent picking date for all distribution and pre-distribution centers. After that, the process ends. The process is uncritical since the bot does not change any data in the system; it only exports data and copies it in Excel files. Since the process is completely rule-based, the exceptions with the picking date are solvable, no unstructured data like, e.g., pictures are used and

all other requirements for RPA are met, the process is selected as the first RPA use case to be automated using UiPath.

The second process is "Availability Maintenance". Figure 3 explains the process flow in the form of a BPMN model.
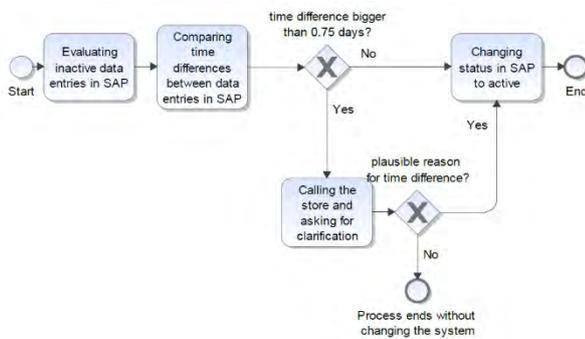

Figure 3: BPMN model "Availability Maintenance", own representation

Each store can change its availability times for delivery whenever needed. Any changes to the availability times cause inactive data entries which need to be changed to active within the process. The process is also executed daily and needs about 10-30 minutes for an employee. The process needs to be executed before 3 PM. The process looks as follows:
An employee logs into SAP at 2:30 PM the latest, calls a transaction, and selects all data with the status "I" for "inactive" in the SAP input mask. After executing, the employee sees a table with all inactive data entries for the relevant stores. The store identification numbers need to be copied and pasted into another transaction. Also, the most recent date indicated needs to be copied for the second transaction.

The next SAP transaction changes the availability entries. The employee enters the store identification numbers, the most recent date from the previous transaction minus one day and executes. The employee sees a table with different data entries. Next, the employee compares every row where the status parameter indicates "inactive" with the previous row to determine the time difference. When the time difference is 0.75 days or less, the employee changes the status, if the difference is greater, the employee calls the store and asks for clarification. Afterward, the process ends.

Consequently, the process is completely rule-based, only includes structured data, is repetitive and stable. The activities which would be implemented with the bot do not contain a system break since all of them happen in SAP. The complexity of the process is low. However, there is still one interesting difference compared to the first process. The "Availability Maintenance" does not only export data but also changes data in SAP. Also, there would be no need for an employee to use SAP anymore for the process. The employee could get all necessary information via email from the bot. Another major reason why RPA is beneficial for this process is that it has to run on the weekends. Stores may change their availability times 24/7, so even on e.g., Saturdays. Since no office employees are available to adjust the status in SAP, the request would have the status "inactive" until Monday. By implementing the process with RPA, this problem would be solved.

Analysis and Redesign
The next two phases in the BPM lifecycle are the analysis and process redesign phases.

The "Gross Load Preview" process has been slightly redesigned before developing the automation. The email which is sent in the "Gross Load Preview" process contains a screenshot of the Excel file. Since screenshots and pictures, in general, are unstructured data and consequently difficult to use with RPA on a table was sent instead of the screenshot.

Another aspect that has been redesigned in the "Gross Load Preview" process is the location of the Excel files. Beforehand, the employees used an internal drive to save their Excel files which meant that only one person at a time could work on a file. Consequently, if the bot aims to write in an Excel file that is locked by another user, the bot will crash. Informing the employees in the department that, e.g., 6-7 AM, no one should access the Excel files was not safe enough. Another solution is to let the bot write in its own Excel files and integrate VLOOKUPs in the department's shared Excel files. One problem with the second solution is that whenever the department's Excel files are changed in their layout, this needs to be done to the bot's files too. Due to bad experiences with the VLOOKUP solution in the past, also this idea was discarded. Since the company plans to switch to SharePoint during 2021 anyway, the traditional drives will not be used anymore. These circumstances led to the decision to create a SharePoint where the bot saves the updated Excel file daily. This can be seen as process optimization.

Also, the "Availability Maintenance" process has been redesigned. The employees need the information which data entries were not changed to "active" since the time difference was too big. This can be implemented by sending an email to the employee with all entries having a larger time difference than 0.75. Emails have not been sent before within the process. Also, the bot changes data in the system. To ensure traceability of the changes, one suggested change is to let the bot write a log with all data which has been changed in SAP.

Implementation
The next step in the BPM lifecycle is the implementation of process automation. In this case, the phase includes primarily the programming with UiPath. For simple click and write activities, low to no coding skills are required.

One example is the following: For both processes, a sequence to log into SAP is easy to implement since it consists of click and write activities only.

For both processes, it is also necessary to develop sequences that require more programming skills. One example is that in both processes, some identification numbers need to be read, copied, and pasted into another transaction. This requires the creation of a temporary data table so that the bot remembers which data to paste into the other transaction.

To implement the logic of changing a data entry status within the "Availability Maintenance" process, also other activities are needed. In this case, the data in the table needs to be read, and values need to be calculated to get the time difference between two rows. Then a so-called "flow decision"-activity helps the bot decide whether it can change the status by itself or should write the information of the entry in an email.

UiPath also has activities that are useful for working with the Orchestrator in the case of non-attended bots. There is an activity called "Get-Credential" which gets the assets that are stored in the Orchestrator. The passwords are encrypted as credentials in the Orchestrator. Through using the activity "Get-Credential", the bot gets the password and username without anyone knowing the password. This solves concerns regarding data security.

The "Gross Load Preview" process was the first process to implement and took about three full days of programming, whereas the "Availability Maintenance" process took only one day of programming. Reasons for the time difference are the lower complexity of the "Availability Maintenance" and more experience with programming in UiPath. Also, some sequences could be reused like the "log into SAP" sequence. Consequently, the initial programming did not take a lot of time.

After the bot development, the bots needed to be tested. The tests showed how volatile the implementation with RPA is. Differing language settings within, e.g., Excel or SAP or different GUI settings in SAP cause bots to crash. However, the goal was to run the bots on virtual machines whose settings would not be changed without an announcement by the Orchestrator administrators. Another problem that came up during the testing was that the bot always crashed if the SAP application was already opened in the background. The solution for this problem was to integrate a sequence at the beginning of each process that closes all applications. Similarly, a sequence was created to close all applications at the end of each process. Those two sequences are useful for any RPA bot and potentially help to make the processes run more stable.

Another example of a bot crash was that MS Teams would open up automatically in the background while the bot was executing the process. The solution was to block any automated program starts in the task manager of the virtual machine. Furthermore, bots crashed because the bot tried to execute activities faster than the application would load the data. A solution for this problem is to integrate "Delay" activities that advise the bot to wait, e.g., two seconds before starting the next activity. Another reason for a bot crash was an empty table in SAP, even though the bot expected the table to include data. This happened when the "Gross Load Preview" bot iterated over the distribution centers, but one distribution center did not pick in the selected time interval. These exceptions were intercepted by adding a sequence that causes the bot to skip the distribution center when the table is empty.

The bot needs about three minutes for the execution of each process. The reason why the bot is not faster than three minutes is that bots need to wait e.g., for transactions in SAP to load just as humans would do since RPA is surface automation only. However, the bot is about 30 minutes faster than manual execution. After two weeks of testing the attended bots and fixing all issues, the bots ran stable.

To let the bots run non-attended on the provided infrastructure different steps need to be considered too. The first step required is to request a service user in the identity management system of the company for the bot.

Next, it is important to assign different Active Directory (AD) groups to the user. These AD groups control which access the user has. The bot user needs the basic package for service users, access to the internet, and access to the virtual machines. Each AD group has a unique name and can be requested through the identity management system of the company.

The next step is to request Office365 rights to ensure that the bot can use, e.g., Excel as well as an SAP user with appropriate SAP roles. Since a lot of damage can be done by mistakenly changing data in the productive system, the SAP user rights management is very strict. The company has an entire department being responsible for user rights management. The idea of assigning user rights to bots was initially seen as a high risk. A solution is to strictly define a role per process. In a meeting between the user rights department, the processes were executed once manually, and a trace was recorded. The trace included the exact description of which transactions were used and if the bot needs read- or also write rights per transaction. Next, the user rights department created the role per process by using a predefined name. The role has been tested by executing the processes once manually with the bot user.

Another step is to set up the virtual machine. This is done after all user rights have been received. The internet connection, SAP GUI settings, and user rights, Office365, SharePoint etc., needs to be set up and tested.

Also, the two bots need to be uploaded to the Orchestrator. Furthermore, the process triggers needed to be set up in the Orchestrator. The "Gross Load Preview" process needs to be executed before 8 AM and the "Availability Maintenance" before 3 PM. Therefore, the daily execution time for the "Gross Load Preview" process is set to 7:30 AM and the time for the "Availability Maintenance" process to 2:30 PM. The time buffer of half an hour is sufficient because the bot execution takes about five minutes only.

Also, the credentials of the bots needed to be set up as assets in the Orchestrator. This is necessary for the bot execution to ensure that the bot has the relevant passwords and usernames. The passwords are encrypted in the Orchestrator. After all the mentioned steps, the bots can run non-attended.

When both bots started running productive, daily check-ups were made for one week to verify whether the data generated by the bot matches the data of the manual process execution. Some further exceptions were found, so the bots needed additional adjustments. One example is that within the "Availability Maintenance" the bot will not send out an email if no data needs to be changed for this day. A sequence has been added in the bot's code so that the logistics department receives an email that informs them that no changes were made.

Monitoring and Control
The last phase in the BPM lifecycle is the monitoring and control phase. The Orchestrator is the tool to monitor and control the non-attended bots. One important tab is the "Automations"-tab. This tab includes the function to trigger the processes. Consequently, the employees can trigger the processes manually. Furthermore, new bots can be uploaded in the Orchestrator, whenever a new RPA case is implemented. Another valuable function within the Orchestrator is the "Jobs"-section, where any member of the folder can see whether the process execution was successful and the screenshots about process execution.

DISCUSSION

The research question was "Which factors should be considered when aiming to introduce RPA in a company or a department of a company?".
Although RPA is sometimes portrayed as a quick fix of the processes, it can be challenging to achieve the business benefits. The case study revealed 20 different lessons learned which can increase the likelihood of a successful RPA introduction in a company:

Lesson Learned 1 - Understanding the technology: Understanding the technology with all its strengths and weaknesses is important to identify suitable use cases and to prevent from seeing RPA as an all-encompassing solution for every process. A lot of research about RPA with the use of various literature must be done by the employees aiming to introduce RPA in a company.

Lesson Learned 2 – Communicate RPA knowledge: The findings about RPA need to be communicated in the company to ensure that the knowledge is accessible for all. The benefits of RPA need to be communicated transparently and understandable. These benefits are, e.g., more time for other tasks, better reports due to fewer mistakes, etc.

Lesson Learned 3 – Define suitable RPA cases: The questionnaire introduced in the "identification" section of the paper provides an orientation which questions may help identifying suitable RPA cases.

Lesson Learned 4 – Start small: Within the case study, two simple processes have been automated. The two processes meet all the requirements for a suitable RPA case. The differences between the processes helped to prove that RPA can be a good solution for different processes. The simplicity of the processes helped that the programming of the bots did not require too much time. Also, the experience which has been generated through the development of the first bot helped to develop the second bot faster.

Lesson Learned 5 - Choose an appropriate RPA provider: The software vendor market should be analyzed to find an RPA vendor which fulfills the specific requirements of the company.

Lesson Learned 6 – Role definition: All involved employees should have defined roles. This means it needs to be documented who is the process expert, who is responsible for the bot development, for the bot users, for the RPA infrastructure, etc.

Lesson Learned 7 - Process analysis and optimization: The added value of automating inefficient processes is a lot smaller than when automating optimized processes. RPA can also be a chance to rethink processes and change them partly to ease automation. Consequently, the introduction of RPA can help to modernize processes in general.

Lesson Learned 8 - Process documentation: A very detailed process documentation is needed e.g., as the basis for the programming. In the given cases, a video where an employee executes the process manually was the best solution since RPA is surface automation. Also, it was useful to record the process execution in a meeting so the questions about the process and its exceptions were clarified as part of the recording. The process documentation is also crucial to have a backup whenever the bot may crash due to unforeseen reasons.

Lesson Learned 9 – Consider User-Bot interaction: The interaction between bots and users should be a seamless as possible. Therefore, e.g., possible file merging conflicts due to bots and employees operating in the same files should be avoided. The usage of SharePoint instead of internal department drives can be a possible solution.

Also, adding a sequence in the bot's code where the bot saves the current file before changing anything can prevent issues with files. Furthermore, seamless interaction can be achieved by exchanging information between bots and employees. In the case study, both bots send emails to an email distribution list in case of successful execution. Additionally, the Orchestrator triggers emails that are sent to the employees whenever a bot does not run successfully. Receiving this email is a sign for the employee to execute the process manually and for the developers a sign to check whether the bot's code needs to be fixed or to find out which other problems could have led to a bot crash.

Lesson Learned 10 – Communication of progress: Continuously communicate the progress and define the next steps in regular appointments to keep up the motivation. Also, regular meetings help to define the next steps and clarify who will fulfill which task until when.

Lesson Learned 11 – Document final approval: After a bot is programmed successfully, get the final approval of the process owners for the bot. In the given cases, this has been done in a meeting, where the attended version of the bot has been shown to the process experts. The approval made by the process owners has been documented.

However, RPA should not be limited to one-time bot implementation but should provide long-term benefits of the RPA program. For that, the following is needed:

Lesson Learned 12 –Involve the department which plans to introduce RPA: This involvement can be eye-opening for a department in terms of what needs to be done to run non-attended bots. One potential advantage could be that inquiries for a bot implementation will be done more consciously. Also, it helps to create an understanding of why the bot implementation may take more time than initially estimated.

Lesson Learned 13 - Proper Orchestrator: Orchestration provides an overview of which bots are scheduled for which time slot, which virtual machine is used for which bot, which bot version is currently used, which bots run successfully, which executions lead to error messages etc. The Orchestrator is an important aspect for the seamless interaction between the bots and the employees since it can be seen as the virtual bot manager.

Lesson Learned 14 - Establish a user rights concept for RPA: What needs to be avoided is, that the bots would have more user rights in systems like SAP than the employees. This would lead to bot super users which could potentially lead to a shadow-IT in a company.

Lesson Learned 15 - Reduce the risk of bots doing mistakes: A solution approach could be to automate non-critical processes only. Also, it may help to have testing systems to ensure that no actual data in the system was changed. For more critical processes other process automation solutions like back-end automation may be a safer solution.

Lesson Learned 16 - Calculation of potential cost savings: The case study's focus was not to generate high savings through introducing RPA since the project goal was to understand the technology RPA, to get familiar with programming in UiPath, understanding how to make processes run non-attended, etc. For the future, economic consideration will become more important since long-term the usage of RPA needs to pay off. It is important to not only consider the programming time effort and compare it to the manual execution time. The case study showed that the effort for programming is only one aspect that consumes time. Getting the necessary user rights, setting up the virtual machines, testing the bots, adjusting them, getting the final approval, uploading the bot in the Orchestrator, updating the assets in the Orchestrator, etc. Additionally, also license costs for the bot users cause additional costs.

Lesson Learned 17 - Understanding that RPA is rather a short to a middle-term solution: RPA can be described as "the first step of the automation journey". However, RPA is also "the quick and dirty solution". RPA is quick since the implementation of bots with UiPath can be done in a few days. It can also be described as "dirty" because it has the potential to cover inefficient processes or outdated IT. To prevent hiding an outdated IT, it may help to assign a lifecycle of six months for each bot before it should be reevaluated.

Lesson Learned 18 – International roll-out of RPA: If a company is operating internationally, the plants in all countries should be either encouraged or even required to harmonize the processes. This leads to maximized time savings generated through RPA. Harmonized processes also pave the way for process automation with other automation solutions. The company analyzed in the case study, currently operates in 14 different countries so if processes are harmonized internationally the time savings created through RPA are larger.

Lesson Learned 19 - Manage the volatility of bots: There are several unforeseen reasons why bots may crash like e.g., MS Teams messages randomly popping up or SAP log-in mask showing up at a different place and the bot could not find the field to type in the correct SAP system. An employee would just intuitively close the pop-up message or drag and drop the SAP window into another position. However, the RPA bots are – at least currently – not able to do so. Also, whenever updates on the virtual machines are made, additional testing and possible adjustments of the bots are needed.

Lesson Learned 20 - Gain employees' trust in the bots: In the first week of running the "Availability Maintenance" process productive, the bot did not send out any emails

since an employee did the process already manually earlier that day. If the employees still do the bot's work manually, time savings do not happen. However, trust in bots may increase over time if the bots run stable and correctly.

Considering these 20 aspects helped in the company of the case study to achieve a successful introduction of RPA. The number of non-attended automations is constantly increasing, new licenses are bought to run more bots simultaneously, and more departments expressed interest in introducing RPA too. However, it needs to be mentioned, that this list of lessons learned focuses on RPA related aspects only. Also, since only one company has been analyzed within the case study, there may be more relevant aspects.

OUTLOOK

RPA can be defined as a "bridge technology" which already has the potential to generate time and money savings for companies but also develops further. Therefore, some developments and trends associated with RPA are explained in the following:

If a company aims to combine RPA with Artificial Intelligence (AI), this step is also called Smart Process Automation (SPA) or Intelligent Process Automation (IPA) (Zhang, 2019). AI provides the possibility to integrate human intelligence in executing tasks, whereas RPA focuses on executing tasks where no to limited human intelligence is required (Zhang, 2019). A combination of both technologies may potentially enable the automation of more complex processes. Consequently, current drawbacks of RPA being only able to automate structured, simple, and repetitive processes could be bypassed in the future. Different RPA providers as, e.g., UiPath are currently already working on adding AI to their RPA solutions (Ribeiro, et. al. 2021).

Another new term in the field of automation is hyperautomation. Hyperautomation integrates different new technologies such as natural language processing (NLP), intelligent optical character recognition (OCR), communication analytics, process optimization, machine learning deployments, and AI into the route of process automation (Walker, 2020). Intelligent OCR potentially helps to read non-structured data from e.g., handwritten scans, which was another limitation of RPA. Furthermore, hyperautomation aims to process high volumes of data seamlessly and automate entire RPA processes in one process (Rauch, 2020).

Other articles predicting the future of RPA mention the possibility of associating RPA with process mining. This could potentially lead to a better chance for successful use of the technology. The benefit would be that companies will not adopt automation for automation's sake, but instead focus on higher success rates (Casey, 2020).

Another potential new development is autonomous automation. The idea behind autonomous automation is that the bots themselves will be enabled to automate processes so human development of bots will not be needed anymore and consequently, even the automation could be automated (Casey, 2020).

To conclude, RPA is a technology with a lot of potentials to generate savings but also with some drawbacks which need to be considered to successfully introduce the technology.

LITERATURE

Bullen, C. V. & Rockart, J. F., (1981). A primer on critical success factors. *Massachusetts Institute of Technology (MIT), Sloan School of Management,* 1220-81.

Casey, K. (2020, 22nd December). *5 Robotic Process Automation (RPA) trends to watch in 2021: How is Robotic Process Automation adoption changing? What about data privacy approaches? Tie-ins to AI tools? Experts share RPA issues to keep on your radar screen.* Retrieved 25th May, 2021 from web address https://enterprisersproject.com/article/2020/12/rpa-robotic-process-automation-trends-watch-2021.

Dumas, M., La Rosa, M., Mendling, J. & Reijers, H. A. (2013). *Fundamentals of Business Process Management.* Berlin, Heidelberg: Springer Berlin Heidelberg.

Flechsig, C., Lohmer, J. & Lasch, R. (2019). Realizing the Full Potential of Robotic Process Automation Through a Combination with BPM. In C. Bierwirth, T. Kirschstein & D. Sackmann (Eds.), *Lecture Notes in Logistics. Logistics Management* (pp. 104–119). Cham: Springer International Publishing.

Gartner. (2020, 27th July). *Gartner Magic Quadrant for Robotic Process Automation.* Retrieved 25th May, 2021 from web address https://www.gartner.com/en/documents/3988021/magic-quadrant-for-robotic-process-automation.

Gartner. (n.d.). Intelligent Business Process Management Suites (iBPMS) Reviews and Ratings. In *Gartner Peer Insights*. Retrieved 25th May, 2021 from web address https://www.gartner.com/reviews/market/intelligent-business-process-management-suites

Koch, C. & Fedtke, S. (2020). *Robotic Process Automation: Ein Leitfaden für Führungskräfte zur erfolgreichen Einführung und Betrieb von Software-Robots im Unternehmen* (1st ed. 2020). Berlin, Heidelberg: Springer Berlin Heidelberg; Imprint: Springer Vieweg.

Langmann, C. & Turi, D. (2020). *Robotic Process Automation (RPA) - Digitalisierung und Automatisierung von Prozessen:*

Voraussetzungen, Funktionsweise und Implementierung am Beispiel des Controllings und Rechnungswesens (1st ed. 2020). Wiesbaden: Springer Fachmedien Wiesbaden; Imprint: Springer Gabler.

Oehlrich, M. (2019). *Wissenschaftliches Arbeiten und Schreiben*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Power, C., van Nueten, N., Chandler, S. & Fulton, M. M. (2017). *Who minds the bots? Why organisations need to consider risks related to Robotic Process Automation.* Retrieved 25th May, 2021 from web address https://www.pwc.com.au/publications/assets/rpa-risk-controls.pdf

Ray, S., Villa, A., Tornbohm, C., Rashid, N. & Alexander, M. (2020, 27th July). *Magic Quadrant for Robotic Process Automation*. Retrieved 19th May, 2021 from web address https://www.gartner.com/doc/reprints?id=1-1ZK435W1&ct=200728&st=sb&mkt_tok=OTk1LVhMVC04ODYAAAF8j8Vnv4QwN8XQXxvw7FkrUjlpzH22dxdOImzPX1sjm6j4lycRI9CuPt7OhE51v1IWqW87jx6mWki0O7vfcKn-IRSVSa_iR3IA4HpZLhIgJA

Rauch, S. (2020, 17th December). *Top Trends in RPA for 2020*. Retrieved 25th May, 2021 from web address https://www.simplilearn.com/top-trends-in-rpa-article.

Ribeiro, J., Lima, R., Eckhardt, T. & Paiva, S. (2021). Robotic Process Automation and Artificial Intelligence in Industry 4.0 – A Literature review. *Procedia Computer Science*, 181, 51–58.

Santos, F., Pereira, R. & Vasconcelos, J. B. (2020). Toward robotic process automation implementation: an end-to-end perspective. *Business Process Management Journal*, 26(2), 405–420.

Trkman, P. (2010). The critical success factors of business process management. International Journal of Information Management, 30(2), 125–134.

Vanwersch, R. J. B., Shahzad, K., Vanderfeesten, I., Vanhaecht, K., Grefen, P., Pintelon, L. & Reijers, H. A. (2016). A Critical Evaluation and Framework of Business Process Improvement Methods. *Business & Information Systems Engineering*, 58(1), 43–53.

Vom Brocke, J. & Rosemann, M. (2010). *Handbook on Business Process Management* 1. Berlin, Heidelberg: Springer Berlin Heidelberg.

Zhang, C. (2019). Intelligent Process Automation in Audit. *Journal of Emerging Technologies in Accounting*, 16(2), 69–88.

Optimierung der Organisationsstrukturen erlaubt modulare Fertigung
Praxisbericht

Guido Piech

ams.Solution AG

Rathausstraße 1, 41564 Kaarst

E-Mail: g.piech@ams-erp.com

ABSTRACT

Im Zuge der Einführung eines Projektmanagement-ERP-Systems gestaltete Purplan seine Prozesse papierlos und digitalisiert und teilautomatisiert seine Fertigung.

Der inhabergeführte Anlagenbauer Purplan konstruiert und fertigt Tanklager, Mischstationen, Reaktions- und Versorgungsanlagen, zu deren Abnehmern international tätige Konzerne sowie Mittelständler aus den Bereichen Chemie, Food, Automotive und Kunststoffverarbeitung zählen. Von der individuellen Projektkonfiguration über die Automatisierungstechnik bis hin zur schlüsselfertigen Übergabe, Inbetriebnahme und Wartung erfolgen alle Leistungen aus einer Hand. Um dieses breite Spektrum möglichst effizient erbringen zu können, entschieden sich die Verantwortlichen für die Implementierung des durchgängigen Projektmanagement-ERP-Systems ams.erp.

Schlüsselwörter: Projektmanagement-ERP-System, ams.erp, Anlagenbauer Purplan

Einleitung

Wenn es einem Unternehmen wie Purplan gelingt, sich nach seiner Gründung 2002 in knapp zwei Jahrzehnten zu einem ca. 200 Mitarbeiter starken Mittelständler mit Niederlassungen in den Vereinigten Staaten und in China sowie entsprechend internationalem Aktionsradius zu entwickeln, benötigt es ab einem gewissen Zeitpunkt eine möglichst strukturierte und effiziente Ablauforganisation. Um diese zu gewährleisten, entschieden sich die Verantwortlichen 2015 nach der ausgiebigen Begutachtung mehrerer ERP-Produkte für die Implementierung von ams.erp von ams.Solution. Ein Hauptgrund war laut Geschäftsführer Andreas Sandmann die Durchgängigkeit des Systems, das er im Vergleich für das gesamtheitlichste hält und in diesem Zusammenhang insbesondere auf das integrierte CRM-Modul verweist. Als weiteres entscheidendes Argument führt er die Funktionsweise der Software an, die speziell die Anforderungen der Losgröße 1+ berücksichtigt.

Das erste Etappenziel: Papierlose Prozess

Die Vorgabe lautete zunächst, die Prozesse papierlos zu gestalten und auf diese Weise verbindliche sowie durchlässige Abläufe zu etablieren. So erhielten beispielsweise die Monteure auf den Baustellen mobilen Zugriff auf das ERP-System, um Stücklisten, Dokumente und Zeichnungen digital einsehen zu können. Sie sollten zudem ihre Personaleinsatzzeiten und Materialentnahmen direkt vor Ort ins System eingeben. Diese Maßnahmen griffen: Papierdokumente fallen im Fertigungs- und Montagebereich inzwischen kaum noch an. War es im Schadensfall zuvor immer sehr aufwendig, aus den Papierdokumenten – so sie denn vollständig waren – die Sachlage zu ermitteln, ist heute zu jedem Zeitpunkt rekonstruierbar, wann welcher Schweißer welche Naht erstellt hat. Über die Stücklistenbuchungen ist zudem bekannt, welche Chargen in den jeweiligen Auftrag eingingen. Alle relevanten Informationen werden in ams.erp anschaulich aufbereitet und stehen baugruppenbezogen „auf Knopfdruck" zur Verfügung, was auch mit Blick auf die steigenden Dokumentationsanforderungen der Kunden nicht unwesentlich ist.

Abb.1: Zu den Abnehmern der Tanklager, Mischstationen, Reaktions- und Versorgungslagern zählen international tätige Konzerne

Auf dem Weg zu einer industriellen Vorproduktion

Der eigentliche Clou besteht darin, dass es mit ams.erp gelungen ist, nicht nur die Digitalisierung ehemals papierbasierter Prozesse voranzutreiben, sondern überdies auch die Art und Weise der Fertigung grundlegend zu verändern. Maßgeblich vorangetrieben wurde das Thema von Konstantin Kröger, der seit 2019

als Head of Digital Transformation auch das ERP-System federführend betreut.

Wie in der Branche üblich, transportierte Purplan die noch unfertigen Komponenten seiner Anlagen über Jahre hinweg zunächst auf die jeweiligen Baustellen und ließ sie vor Ort für die Montage vorbereiten. „Wir stießen dann darauf, dass es erheblich einfacher ist, die Anlagen nicht mehr wie zuvor als monolithischen Gesamtblock zu betrachten, sondern mithilfe des ERP-Systems stücklistenbasiert in kleinere und kleinste Einheiten herunterzubrechen", berichtet Konstantin Kröger. Eine Erkenntnis mit großer Wirkung, denn so lässt sich nicht nur jede einzelne Rohrleitung exakt vorplanen und erfassen. Vielmehr können die Teile nun allesamt auftrags- und projektbezogen vor dem Transport auf dem eigenen Werksgelände gefertigt werden und müssen beim Kunden letztlich ‚nur noch‘ mit wenigen Schweißnähten zusammengefügt werden. Dadurch kann Purplan in diesem rein handwerklichen Bereich mit ausländischen Billiganbietern konkurrieren und zugleich eine höhere Qualität garantieren.

Im Rahmen der Vorplanung wird der künftige Standort der Anlagen mit einem Hochleistungslaserscanner vermessen. Auf Basis des virtuellen 3-D-Modells erfolgt anschließend die Anlagenkonstruktion im CAD-System Autodesk Plant-3D. Alle relevanten Daten werden im ERP-System erfasst, über das schließlich die Produktion und Montage der gesamten Anlage gesteuert wird. Zugleich entsteht ein digitaler Zwilling, der eine virtuelle Begehung sowie eine Simulation der Anlagen ermöglicht. Wünsche, Änderungen und Anregungen von Kunden können somit bereits vor dem Produktionsstart aufgenommen werden, was die Gefahr von Fehlentwicklungen immens reduziert.

Die neuen Strukturen zeigen sich in den Werkshallen. Inzwischen existieren festinstallierte Fertigungsplätze mit Absaugvorrichtungen, Schweißgeräten sowie Terminals und Bildschirmen, auf denen die Rohrleitungen dargestellt sind und über die die Monteure und Schweißer ihre Arbeitsgänge ins System melden. Viele Arbeitsgänge erfolgen modular bzw. teilautomatisiert. Laut Andreas Sandmann entfernt sich sein Unternehmen damit zunehmend von seinem handwerklichen Ursprung in Richtung einer planerischen, ingenieurstechnischen Prägung. Für Konstantin Kröger der größte Effekt, der sich aus der täglichen Arbeit mit dem ERP-System ergibt: „Ohne ams.erp wäre diese Entwicklung meiner Ansicht nach unmöglich gewesen."

Nahe am Softwarestandard

Ein großer Vorteil ist, dass diese weitreichenden organisatorischen Verbesserungen umgesetzt werden konnten, ohne das ERP-System zu verbiegen. Es gibt zwar einige Anpassungen, doch generell bewegt sich Purplan sehr nahe am Softwarestandard. Dies war auch von Beginn an beabsichtigt, denn auf diese Weise

profitiert das Unternehmen zum einen von den bewährten Best-Practice-Prozessen, zum anderen verlaufen Releasewechsel unkompliziert.

Bei den Niedersachsen sind neben dem Kernsystem viele Komponenten aus dem ams-Produktportfolio im Einsatz – angefangen bei ams.finance über das integrierte Dokumentenmanagement, die strategische Planung und ams.mobile zur Personalzeit- und Materialerfassung bis hin zum ebenfalls integrierten Variantenkonfigurator. Die Durchgängigkeit des Datenflusses sorgt dafür, dass die Prozessketten ineinandergreifen und mögliche Fehlerquellen von vornherein minimiert werden.

Abb. 2: 75% der Montagearbeiten finden mittlerweile in den Werkshallen statt.

Kürzere Durchlaufzeiten

Den Variantenkonfigurator nutzt Konstantin Kröger, „um die Struktur im Auftrags- und Angebotswesen grundlegend zu verbessern". Mit Sollkostenabfragen und dem Vergleich zwischen Soll und Budget liefert ams.erp eine sehr präzise Kostenkalkulation und sorgt für Planungssicherheit. „Die Durchlaufzeiten in der Projektabwicklung haben sich seit dem Einsatz der Software in jedem Fall signifikant verringert", versichert der Prozess- und IT-Experte und verweist darauf, dass der Anspruch an das Projektmanagement vonseiten der Kunden insgesamt immens gewachsen sei. Er vertritt sogar die These, ohne ERP-System könne Purplan unter wirtschaftlichen Gesichtspunkten wahrscheinlich gar keine Aufträge mehr kostendeckend durchführen.

ams.Solution AG: Die Welt für Projektmanagement-ERP

Das Beratungs- und Softwarehaus ams.Solution AG, ein Unternehmen der ams.Group, ist auf die Projektmanagement-Anforderungen von Einzel-, Auftrags- und Variantenfertigern spezialisiert. Seit mehr als 30 Jahren werden auf Basis der branchenorientierten Business-Software ams.erp schlanke und dynamische Unternehmensprozesse entlang der gesamten Wertschöpfungskette realisiert. Die Gruppe verfügt über das Know-how aus europaweit mehr als 1.000 erfolgreichen ERP-Implementierungen im Maschinen-, Anlagen- und Apparatebau, im Verpackungsmaschinen-

bau, im Werkzeug- und Formenbau, im Stahl-, Metall- und Komplettbau, im Schiffbau und in der maritimen Industrie sowie im Laden- und Innenausbau, im Sonderfahrzeugbau und in der Lohnfertigung.

ams.erp bildet die zentrale Datendrehscheibe für sämtliche Unternehmensprozesse. Wesentliche Aufgabe ist das Zusammenführen innovativer Digitaltechnologien und klassischer, meist mechanischer und elektronischer Steuerungseinrichtungen. Hierzu zählt auch die standortübergreifende Vernetzung – vertikal in die Fertigung hinein und horizontal über Systeme und Firmengrenzen hinweg. Die flexible Business-Software erhöht die Planungssicherheit, Profitabilität und Wettbewerbsfähigkeit von Unternehmen mit komplexem Projektgeschäft.

Entwicklung eines Kriterienkatalogs zur Gegenüberstellung von Business-Intelligence-Lösungen in der Cloud am Beispiel von Snowflake und der SAP Data Warehouse Cloud

Alexander Dörr

Technische Hochschule Mittel-
hessen

Fachbereich MND
Wilhelm-Leuschner-Str. 13
61169 Friedberg
alexander.doerr@mnd.thm.de

Prof. Dr. Harald Ritz

Technische Hochschule Mittel-
hessen

Fachbereich MNI
Wiesenstraße 14
35390 Gießen
harald.ritz@mni.thm.de

Martin Ried (M. Sc.)

INFOMOTION GmbH

BU Wahl
Westhafenplatz 1
60327 Frankfurt am Main
martin.ried@infomotion.de

Kategorie

Masterarbeit

Schlüsselwörter

Business Intelligence, Cloud-Computing, Softwareeva-
luation, Nutzwertanalyse, SaaS, Kriterienkatalog, SAP
Data Warehouse Cloud, Snowflake

Zusammenfassung

Das schnelle Treffen fundierter Entscheidungen ist im
Zeitalter der digitalen Transformation für Unternehmen
ein entscheidender Wettbewerbsfaktor, um frühzeitig auf
Änderungen im Marktumfeld zu reagieren und somit ei-
nen Wettbewerbsvorteil gegenüber der Konkurrenz zu
erzielen. Zur Entscheidungsunterstützung werden in
der Regel *Business-Intelligence-und-Analytics*-Systeme
(BIA-Systeme) für die Analyse großer Datenmengen ein-
gesetzt. Diese lassen sich nicht nur *On-Premises*, sondern
auch in der Cloud betreiben, sodass zentrale Merkmale
des Cloud-Computings wie die Skalierbarkeit zu einer
optimierten Entscheidungsfindung beitragen.

Eine Herausforderung für Unternehmen ist die adäquate
Auswahl einer Cloud-BIA-Lösung, da der Reifegrad der
am Markt verfügbaren Alternativen hinsichtlich des
spezifischen Anwendungsfalls und unter Berück-
sichtigung aktueller Anforderungen an eine moderne
Analyselandschaft zu evaluieren ist. Daraus resultierend
ist das Untersuchungsziel die Konzeption eines
herstellerunabhängigen Kriterienkataloges zur fundierten
Auswahl einer Cloud-BIA-Lösung. Am Beispiel der
Gegenüberstellung der SaaS-Produkte *Snowflake* und
SAP Data Warehouse Cloud (SAP DWC) ist zu eruieren,
welche Lösung im Hinblick auf die in aktuellen
Publikationen benannten Anforderungen am besten
geeignet ist und für welche Anwendungsszenarien sie
jeweils prädestiniert sind.

Zur Ermittlung relevanter Kriterien wurden Publikatio-
nen mit einer Softwareauswahlkomponente im Cloud-
oder BIA-Umfeld gesichtet. Dabei konnte festgestellt
werden, dass Kosten, Sicherheit und Performance die
wichtigsten Einflussfaktoren bei der Auswahl einer
Cloud-BIA-Lösung sind. Die gewonnenen Erkenntnisse
sind in die Konzeption des Kriterienkatalogs eingeflos-
sen, der aus 32 Kriterien besteht. Diese sind in die neun
Kategorien Interoperabilität, Data Governance, Soft-
wareergonomie, Analyse, PaaS, Support, Cloud-Dienst-
leister & Kosten, Sicherheit und Architektur unterteilt.

Darüber hinaus wurde eine prototypische Data-Wa-
rehouse-Implementierung im Kontext eines fiktiven In-
dustrieunternehmens durchgeführt, wodurch sowohl für
Snowflake als auch für die DWC praktische Beispiele zur
Bewertung einzelner Kriterien herangezogen werden
konnten. Die Evaluation der Lösungen ist anhand einer
Nutzwertanalyse erfolgt, die neben einem Gewichtungs-
szenario, das auf den Erkenntnissen der Literaturstudie
basiert, auch ein zweites Szenario berücksichtigt, das
Kriterien im Bereich *Self-Service-BI* (SSBI) stärker ge-
wichtet.

Im Zuge der Evaluation hat sich zum Betrachtungszeit-
punkt herausgestellt, dass Snowflake bezüglich der An-
forderungen an eine moderne Analyseplattform in den
Kategorien Sicherheit, Interoperabilität, Support, PaaS
und Cloud-Dienstleister & Kosten besser qualifiziert ist
als die DWC. Dennoch weist Snowflake Defizite bei in-
tegrierten Visualisierungsfunktionalitäten und dem Auf-
bau eines Data Lakes auf. Dementsprechend ist
Snowflake für den Einsatz als Data Warehouse geeignet,
das Visualisierungs- und Analyselösungen mit Daten
versorgt und Datenbestände aus Quellsystemen konsoli-
diert.

Die DWC ist im Hinblick auf Anforderungen des Be-
reichs SSBI in den Kategorien Architektur, Analyse,
Data Governance und Softwareergonomie für die Ver-
wendung durch Fachabteilungsmitarbeiter geeigneter als
Snowflake. Jedoch erschweren funktionale Einschrän-
kungen den alleinigen Einsatz der Lösung zum Aufbau
einer modernen Analyseplattform. Trotzdem kann die
DWC als zusätzliche Schicht zwischen Datenbereitstel-
lung und Visualisierung Mehrwerte liefern, um eine da-
tengetriebene Unternehmenskultur voran zu treiben.

Analyse und Implementierung einer Depotverwaltung auf der Salesforce-Plattform

Max Julius Euler

Fachbereich MNI

Technische Hochschule
Mittelhessen
Wiesenstraße 14
35390 Gießen
max.julius.euler@mni.thm.de

Frank Kammer

Fachbereich MNI

Technische Hochschule
Mittelhessen
Wiesenstraße 14
35390 Gießen
frank.kammer@mni.thm.de

Harald Ritz

Fachbereich MNI

Technische Hochschule
Mittelhessen
Wiesenstraße 14
35390 Gießen
harald.ritz@mni.thm.de

Kategorie

Bachelorarbeit

Schlüsselwörter

Business Intelligence, Salesforce-Plattform

Salesforce Depotverwaltung

Die Erstellung einer umfangreichen und lückenlosen Darstellung der Kunden eines Finanzunternehmens ist eine der großen wirtschaftlichen Herausforderungen der letzten Jahre, welche auch durch die fortschreitende Digitalisierung der Unternehmensprozesse an Bedeutung gewonnen hat. Unternehmen aus der Finanzbranche sind dafür bekannt, neue Technologien, speziell im Hinblick auf das sogenannte Cloud-Computing, frühzeitig zu übernehmen. Durch den Einsatz verschiedenster Systeme und der Gewinnung von Daten aus unterschiedlichsten Quellen, entsteht eine große Komplexität in der Zusammenführung dieser Daten. Eine Arbeit ist effizienter und einfacher, wenn Kundendaten an einem Ort gesammelt, abgerufen und bearbeitet werden können. Diese Möglichkeit bietet u.a. die Salesforce Plattform, welche ihre sogenannten CRM-Funktionalitäten, was für Customer Relationship Management steht, als Plattform zur Verfügung stellt.

Auf der technischen Seite hat die Arbeit gezeigt, dass das Einspielen, Pflegen, Berechnen und Darstellen aller benötigten Daten und Informationen auf der Salesforce-Plattform umsetzbar ist. Es wurden verschiedene Technologien und Konzepte der Plattform besprochen und ihre Vorteile für die konkreten Anforderungen verdeutlicht. Durch die Verwendungvon APEX REST Klassen können Daten empfangen und in das erstellte Datenmodell eingespielt werden. Der Einsatz von Batch Klassen in Kombination mit dem Schedulable Interface ermöglicht eine automatische und regelmäßige Verarbeitung großer Datenmengen, um daraus Informationen, wie die Performance Kennzahlen oder den Depotverlauf, zugenerieren. Auch die sinnvolle grafische Darstellung der Daten kann mithilfe von Berichten und Diagrammen erreicht werden. Dies wurde bei der Implementierung der Datensatzseiten gezeigt. Sowohl das Donut-Diagramm zur Darstellung der Aufteilung des Depots, als auch das Linien-Diagramm zur Darstellung des Depotverlaufs bieten wertvolle Informationen und eine benutzerfreundliche Darstellung für die zuständigen Mitarbeiter. Es hat sich also herausgestellt, dass die Salesforce-Plattform durchaus in der Lage ist, eine Depotverwaltung, unter gewissen Einschränkungen, zu betreiben. Diese Einschränkungen beziehen sich auf den möglichen Vertrieb der Applikation, wobei hier speziell die Abhängigkeit von einem Drittanbieter zur Bereitstellung der benötigten Daten zu nennen ist. Es ist eine Applikation entstanden, welche entscheidende Informationen zu allen gewünschten Aktien ins System einspielen, dort pflegen und darstellen kann. Es werden Performance Kennzahlen berechnet, die ausschlaggebend für die Bewertung der Aktien sein können. Die Depots werden aus skalierbaren Datensätzen aufgebaut. Alle entscheidenden Transaktionen können festhalten werden. Zudem liefert das Datenmodell eine Grundlage für eine beliebige Weiterverarbeitung der gewonnen Informationen.
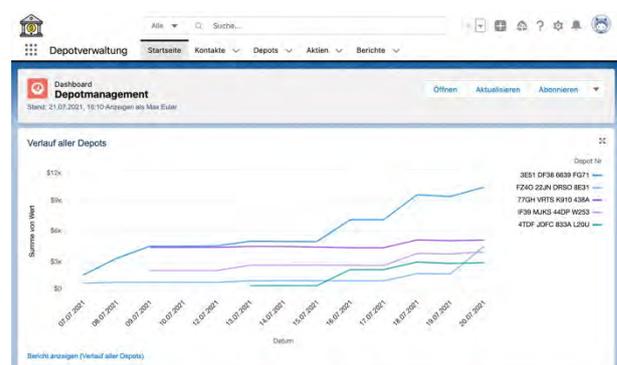
LITERATUR

Reinheimer, Stefan: Cloud Computing: Die Infrastruktur der Digitalisierung, Springer Fachmedien Wiesbaden GmbH, Wiesbaden, 2018, https://doi.org/10.1007/978-3-658-20967-4

Nicoletti, Bernardo, Cloud Computing in Financial Services, Palgrave Macmillan UK, 2013, https://doi.org/10.1057/9781137273642

Performance-Nutzen-Analyse: Automatisierung eines Data Warehouse

Untersuchung am Beispiel des Datavault Builder auf verschiedenen Zielplattformen in der Azure Cloud

Kevin Ullrich

Technische Hochschule
Mittelhessen

Fachbereich MND
Wilhelm-Leuschner-Straße 13
61169 Friedberg
E-Mail:
kevin.ullrich@mnd.thm.de

Prof. Dr. Harald Ritz

Technische Hochschule
Mittelhessen

Fachbereich MNI
Wiesenstraße 14
35390 Gießen
E-Mail:
harald.ritz@mni.thm.de

Arthur Arendt

OPITZ CONSULTING
Deutschland GmbH

Am Weidenring 56
61352 Bad Homburg v. d. H.
E-Mail:
arthur.arendt@opitz-consulting.com

Kategorie

Masterarbeit

Zusammenfassung

In der vorliegenden Arbeit wird eine Performance-Nutzen-Analyse vorgenommen, untersucht am Beispiel des Datavault Builder zur automatischen Data Warehouse-Erstellung auf verschiedenen Zielplattformen (konkret: Azure SQL-Datenbank, Exasol und Snowflake) in der Azure Cloud. Die Aufgabenstellung im Zusammenhang mit der vorliegenden Arbeit besteht im Vergleich verschiedener Zielplattformen in Kombination mit dem Datavault Builder und darin, eine Aussage über die Vorteilhaftigkeit einer bestimmten Zielplattform für den Datavault Builder im Hinblick auf die Performance und Nutzung geben zu können.

Zudem soll die Kenntnis der Arbeit es ermöglichen, die Thematik der Data Warehouse-Automatisierung tiefer zu durchdringen, um die jeweils bestmögliche Nutzung in Kundenprojekten zu garantieren. Dazu wird ein Benchmark in unterschiedlich großen Skalierungsausprägungen verwendet. Qualitative Aspekte werden mit Hilfe einer Nutzwertanalyse bewertet.

Als Ergebnis wird deutlich, dass sich die Datenbank Exasol für alle drei TPC-H-Datensatzgrößen trotz des schlechteren Nutzwertes als vorteilhafter herausgestellt hat: zum einem deshalb, weil sie in jeder Skalierungsstufe die geringste Gesamtlaufzeit gegenüber der Azure SQL-Datenbank und Snowflake erzielt hat; zum anderen aus dem Grund, weil sie nur einen Bruchteil der Kosten der anderen beiden Zieldatenbanken verursacht.

Die Azure SQL-Datenbank ist aufgrund ihrer langen Laufzeiten und den damit verbundenen Kosten weniger für den Einsatz mit dem Datavault Builder zu empfehlen.

Snowflake hingegen ist zwar nicht viel langsamer als Exasol, verursacht aber erheblich höhere Kosten. Der Vorteil, den Snowflake jedoch gegenüber Exasol hat, ist darin zu sehen, dass Snowflake als Software-as-a-Service-Lösung bereitgestellt wird, bei der Snowflake alle Aspekte der Verwaltung abdeckt.