

Projektträger für das



Schlussbericht vom 09.03.2026

zum IGF-Vorhaben 01|F23010N

Thema

Operative Produktionsprogrammplanung durch selbstoptimierende Materialflusssimulationen mittels Reinforcement Learning (PrOPPlan)

Berichtszeitraum

01.07.2023 bis 31.12.2025

Forschungsvereinigung

Bundesvereinigung Logistik (BVL) e.V.

Schlachte 31

28195 Bremen

Forschungseinrichtung(en)

RIF Institut für Forschung und Transfer e.V.

Joseph-von-Fraunhofer-Str. 20

44227 Dortmund

Inhaltsverzeichnis

Abbildungsverzeichnis.....	3
Tabellenverzeichnis.....	4
1 Einleitung	5
1.1 Problemstellung	5
1.2 Zielsetzung und Arbeitshypothesen.....	5
2 Stand der Technik zu Projektstart.....	6
2.1 Grundlagen der operativen Produktionsprogrammplanung	6
2.2 Simulation in Materialflusssystemen	7
2.3 Anwendung von Reinforcement Learning zur selbstlernenden Simulationsoptimierung.....	9
3 Durchgeführte Arbeiten und Ergebnisse	11
3.1 Arbeitspaket 1: Ermittlung der PA-Anforderungen und -Bedarfe für Simulationen in Materialflusssystemen	12
3.2 Arbeitspaket 2: Ableitung der relevanten Produktionsparameter für selbstlernende Simulationsumgebungen.....	14
3.3 Arbeitspaket 3: Entwicklung der Methodik zur selbstoptimierenden Simulation mittels RL	16
3.4 Arbeitspaket 4: Anwendung und Optimierung der entwickelten Methodik in realen Materialflusssystemen	19
3.5 Arbeitspaket 5: Überführung der Methodik in ein IT-Werkzeug und vorhabenbegleitende Optimierung	21
3.6 Arbeitspaket 6: Projektmanagement, Öffentlichkeitsarbeit, Dokumentation und Ergebnistransfer	26
4 Verwendung der Zuwendung	27
5 Notwendigkeit und Angemessenheit der geleisteten Arbeit.....	28
6 Darstellung des wissenschaftlich-technischen und wirtschaftlichen Nutzens der erzielten Ergebnisse insbesondere für KMU sowie ihres innovativen Beitrags und ihrer industriellen Anwendungsmöglichkeiten	29
7 Wissenstransfer in die Wirtschaft.....	31
8 Durchgeführte Transfermaßnahmen	32
9 Geplante spezifische Transfermaßnahmen nach der Projektlaufzeit	33
10 Literatur	34

Abbildungsverzeichnis

Abbildung 1: Kernaufgaben der Produktionsplanung und -steuerung gemäß Aachener PPS-Modell (Schuh et al. 2012, S. 31)	7
Abbildung 2: Darstellung von Änderungsdynamik und Planungshorizont von produktionstechnischen Aufgaben der simulationsunterstützten Produktionsplanung (Gierth et al. 2006, S. 665) und Verortung des Wirkungsbereich von PrOPPlan.....	8
Abbildung 3: Schematisches Wirkprinzip beim Reinforcement Learning (Sutton et al. 2018, S. 38; West et al. 2021, S. 348)	9
Abbildung 5: Mathematische Formalisierung der Produktionssysteme erforderlich für Reinforcement Learning	14
Abbildung 6: Arbeitsstation als zentrales Element von Produktionssystemen.....	15
Abbildung 7: Tab 1 der Benutzeroberfläche von Propplan – Eingabe von Produktionsressourcen	22
Abbildung 8: Eingabe von Produktvorranggraphen inklusive Operationsvoraussetzungen, Bearbeitungsdauer und Produkt	23
Abbildung 9: Simulationsdarstellung eines realen Produktionssystems mit dem Arbeitsplan, welcher die kürzeste Durchlaufzeit hat	24
Abbildung 10: Konfiguration von RL-MuZero.....	25
Abbildung 11: Konfiguration des Beobachtungsraums.....	25
Abbildung 12: Konfiguration des Aktionsraums.....	25
Abbildung 13: Konfiguration der Belohnungsfunktion.....	25

Tabellenverzeichnis

Tabelle 1: Eingesetzte personelle Ressourcen im Berichtszeitraum	27
Tabelle 2: Durchgeführte Transfermaßnahmen während der Projektlaufzeit.....	32
Tabelle 3: Geplante Transfermaßnahmen nach Projektende	33

1 Einleitung

Die Produktionsplanung und -steuerung (PPS) stellt einen zentralen Bestandteil industrieller Wertschöpfungs-systeme dar. Insbesondere die operative Produktionsprogrammplanung umfasst kurzfristige Entscheidungen über die mengen- und zeitmäßige Zusammensetzung des Produktionsprogramms und beeinflusst damit unmittelbar die Leistungsfähigkeit von Produktionssystemen. Typische Planungsaufgaben in diesem Kontext sind unter anderem die Reihenfolgenplanung, die Losgrößenbildung, die Feinterminierung sowie die Auftragsfreigabe. In der industriellen Praxis werden diese Entscheidungen häufig sequenziell getroffen und basieren vielfach auf Erfahrungswissen der verantwortlichen Planerinnen und Planer.

Gleichzeitig stehen produzierende Unternehmen zunehmend vor der Herausforderung, ihre Produktionssysteme unter dynamischen Rahmenbedingungen effizient zu betreiben. Variierende Nachfrage, steigende Variantenvielfalt sowie kurzfristige Störungen im Produktionsablauf erfordern eine hohe Anpassungsfähigkeit der operativen Planung. Klassische deterministische Planungsverfahren können die komplexen Wechselwirkungen innerhalb eines Produktionssystems jedoch nur eingeschränkt berücksichtigen.

Vor diesem Hintergrund gewinnen simulationsgestützte und datengetriebene Ansätze zur Entscheidungsunterstützung zunehmend an Bedeutung. Insbesondere die Kombination von ereignisdiskreter Materialflusssimulation mit Methoden des maschinellen Lernens eröffnet neue Möglichkeiten zur automatisierten Generierung von Planungsempfehlungen. Das Forschungsvorhaben PrOPPlan griff diesen Ansatz auf und untersuchte die Entwicklung einer Methodik zur selbstoptimierenden Produktionsprogrammplanung. Im Projekt wurde dazu eine Methodik entwickelt, die Materialflusssimulation mit Reinforcement Learning kombiniert und mehrere zusammenhängende Planungsaufgaben der operativen PPS simultan berücksichtigt.

1.1 Problemstellung

In der industriellen Praxis erfolgt die operative Produktionsprogrammplanung häufig in mehreren aufeinander aufbauenden Planungsschritten, beispielsweise durch getrennte Verfahren zur Losgrößenbildung, Reihenfolgenplanung und Terminierung. Diese schrittweise Vorgehensweise führt dazu, dass Annahmen aus vorgelagerten Planungsschritten später nur eingeschränkt korrigiert werden können und somit häufig lediglich lokal optimale Planungslösungen entstehen. Gleichzeitig erfordern zunehmende Dynamiken in Produktionssystemen – etwa durch variierende Auftragsstrukturen oder kurzfristige Störungen – häufige Neu- oder Anpassungsplanungen, die mit erheblichem Aufwand verbunden sind.

Für kleine und mittlere Unternehmen (KMU) stellt dies eine besondere Herausforderung dar, da operative Planungsentscheidungen häufig stark auf Erfahrungswissen einzelner Mitarbeitender angewiesen sind und gleichzeitig nur begrenzte Ressourcen für komplexe Planungsverfahren oder spezialisierte Softwarelösungen zur Verfügung stehen.

1.2 Zielsetzung und Arbeitshypothesen

Ziel des Forschungsvorhabens PrOPPlan war die Entwicklung einer Methodik zur datengetriebenen Unterstützung der operativen Produktionsprogrammplanung durch die Kombination von Materialflusssimulation und Reinforcement Learning. Im Rahmen des Projekts wurde untersucht, inwieweit mehrere zusammenhängende Planungsaufgaben – insbesondere Reihenfolgenplanung, Losgrößenbestimmung, Feinterminierung

und Auftragsfreigabe – simultan betrachtet und mithilfe selbstlernender Verfahren optimiert werden können.

Auf Grundlage des Standes der Forschung sowie der praktischen Anforderungen der beteiligten Unternehmen wurden Arbeitshypothesen formuliert. Diese gingen davon aus, dass durch die Integration selbstlernender Verfahren in simulationsbasierte Planungsmodelle eine Verbesserung der Planungsgüte sowie eine Reduktion des Planungsaufwands erreicht werden kann. Darüber hinaus wurde angenommen, dass eine geeignete Gestaltung der Methodik sowie deren prototypische Umsetzung in einem nutzerorientierten IT-Werkzeug die praktische Anwendbarkeit für kleine und mittlere Unternehmen ermöglicht. Die Verifizierung bzw. Falsifizierung dieser Arbeitshypothesen erfolgt im weiteren Verlauf des Vorhabens anhand der entwickelten Methodik, der prototypischen Umsetzung sowie der Anwendung in den ausgewählten Anwendungsfällen.

2 Stand der Technik zu Projektstart

2.1 Grundlagen der operativen Produktionsprogrammplanung

Die Produktionsplanung und -steuerung (PPS) bildet den Kern aller Industrieunternehmen (Günther et al. 2016, S. 8; Vahrenkamp et al. 2008, S. 110ff.; Wiendahl 2014, S. 278). Dabei nimmt die in diesem Forschungsvorhaben betrachtete *operative* PPP, neben Aufgaben wie Mengenplanung, Termin- und Kapazitätsplanung, Auftragsveranlassung und -freigabe oder Auftragsüberwachung, eine Hauptfunktion der PPS ein (Schlüter 2001, S. 12f.; Schuh et al. 2012, S. 30). Betriebliche Aufgabenstellungen der PPP werden i. d. R. anhand von drei zeitlichen Ebenen in strategisch, taktisch und operativ unterschieden (Corsten et al. 2016, S. 26; Dangelmaier 2009, S. 9). Während die *strategische* PPP unternehmerisch-langfristige Grundsatzentscheidungen umfasst, wie die Festlegung von Produkt- und Geschäftsfeld sowie vom (geografischen) Marktgebiet (Jacob 1996, S. 1469), beinhaltet die *taktische* PPP die mittelfristige Konzeption und Entwicklung neuer Varianten bekannter Produkte sowie neuer Anwendungsgebiete vorhandener Produkte in gegebenen Produktfeldern (Jung 2016, S. 467). Im Fokus des Vorhabens steht durch die *operative* PPP insb. die Produktionsprogrammerstellung, also sämtliche kurzfristigen Anpassungsentscheidungen, der mengen- und qualitätsmäßigen Zusammensetzung des Produktionsprogramms (Günther et al. 2016, S. 4; Jung 2016, S. 468). Unternehmen legen im Produktionsprogramm fest, welche Produktarten in welchen Mengen und in welchen Planungsperioden hergestellt werden sollen bzw. können (Jacob 1996, S. 1468). Außerdem enthält das Produktionsprogramm Angaben über die zeitliche Verteilung von Produktion und Absatz innerhalb dieses Planungszeitraums (Hoschka 1963, S. 20). Zu diesen Tätigkeiten zählen die Aufgaben der Eigenfertigungsplanung im Sinne des *Aachener PPS-Modells* (Schuh et al. 2012, S. 30), wie bspw. die Losgrößenrechnung, die Ressourcenfeinterminierung oder die Reihenfolgeplanung (s. Abbildung 1). Um begrifflichen Missverständnissen vorzubeugen, sei darauf hingewiesen, dass die in der Abbildung als „Produktionsprogrammplanung“ bezeichneten Aufgaben im Aachener Modell lediglich strategische Tätigkeiten bezeichnen und nicht von der operativen PPP eingeschlossen werden (Schuh et al. 2012, S. 33). Zusammenfassend schlägt die operative PPP, ebenso wie ProPPlan, eine organisatorische Brücke zwischen allen kurzfristigen Planungstätigkeiten eines Unternehmens (Hoschka 1963, S. 20), die in direktem Zusammenhang zum Produktionsprogramm stehen. Das Hauptziel des Vorhabens liegt in der kohärenten Lösung dieser operativen PPP-Aufgaben.

Produktions- programmplanung	Produktions- bedarfsplanung	Eigenfertigungsplanung und -steuerung	Fremdbezugsplanung und -steuerung
<ul style="list-style-type: none"> ▪ Absatzplanung ▪ Primärbedarfsplanung ▪ Ressourcengrobplanung 	<ul style="list-style-type: none"> ▪ Brutto- und Nettosekundär- bedarfsermittlung ▪ Beschaffungsartzuordnung ▪ Durchlaufterminierung ▪ Kapazitätsbedarfsermittlung ▪ Kapazitätsabstimmung 	<ul style="list-style-type: none"> ▪ Losgrößenrechnung ▪ Feinterminierung ▪ Ressourcenfeinplanung ▪ Reihenfolgeplanung ▪ Verfügbarkeitsprüfung ▪ Auftragsfreigabe 	<ul style="list-style-type: none"> ▪ Bestellrechnung ▪ Angebotseinholung und -bewertung ▪ Lieferantenauswahl ▪ Bestellfreigabe

Abbildung 1: Kernaufgaben der Produktionsplanung und -steuerung gemäß Aachener PPS-Modell (Schuh et al. 2012, S. 31)

Wie in Abbildung 1 hervorgehoben konzentrieren sich die Optimierungsempfehlungen auf Tätigkeiten der Eigenfertigungsplanung und -steuerung. Die nachfolgende Auflistung erläutert diese Tätigkeiten (Schuh et al. 2012, S. 53ff.):

- **Losgrößenrechnung:** Bildung von wirtschaftlich optimalen Losen, in denen Teillose aus unterschiedlichen Kundenaufträgen zusammengefasst werden, und die intuitiv oder formelbasiert sowie einmalig oder wiederkehrend erfolgt
- **Feinterminierung:** Ermittlung von Start- und Endterminen der Arbeitsgänge für die gebildeten Fertigungslose
- **Ressourcenfeinplanung:** Gegenüberstellung von Kapazitätsbedarf und -angebot auf Basis der Feinterminierung
- **Reihenfolgeplanung:** Bildung einer optimalen Abarbeitungsreihenfolge für vorgesehene Arbeitsvorgänge, die oft nach festen Selektionskriterien (z.B. Prioritätsregeln) oder Kumulationskriterien (z.B. Rüstzeitminimierung) erfolgt
- **Verfügbarkeitsprüfung:** Steuernde Aufgabe zum Abgleich erforderlicher Kapazitäten (u.a. Material o. Kapazitäten)
- **Auftragsfreigabe:** Veranlassung der Ressourcenbereitstellung mit entsprechenden Freigaberegeln oder -verfahren

Die aufgelisteten Tätigkeiten basieren auf konkurrierenden Zielsetzungen, weshalb eine konsekutive Vorgehensweise anfällig für die Identifikation lokal-optimaler Gesamtlösungen aufgrund von Annahmefehlern vorangegangener Tätigkeiten ist (Phansalkar et al. 1995, S. 950ff.; Seyyedabbasi et al. 2021, S. 2ff.). Dagegen strebt das Vorhaben eine Kombination von mehreren Planungsaspekten an, durch die die Zielsetzung einzelner Planungsschritte integriert und eine global-optimale Planungsempfehlung erstellt werden kann. Der Fokus des Vorhabens liegt auf einer selbstoptimierten Reihenfolgenplanung mit gleichzeitiger Bestimmung idealer Losgrößen, integrierter Feinterminierung und anschließender Auftragsfreigabe. Die noch verbleibenden Tätigkeiten der Eigenfertigungsplanung und -steuerung, also Ressourcenfeinplanung und Verfügbarkeitsprüfung, betrachten wir als sekundäre Optimierungstätigkeiten. Sie setzen einen zusätzlichen Kapazitätsvergleich mit Realverfügbarkeiten voraus und werden nach Bedarf und Möglichkeit im Vorhaben berücksichtigt. Alle bisherigen Arbeiten zur datengetriebenen Optimierung konzentrieren sich primär auf eine Planungsaufgabe, bspw. auf die Reihenfolgenplanung (Tassel et al. 2021, S. 3; Waschneck et al. 2018, S. 1268f.) oder auf steuernde Funktionen in der Auftragsfreigabe (Heng et al. 2004, S. 1418ff.; Kuhnle et al. 2021, S. 857ff.).

2.2 Simulation in Materialflusssystemen

Prinzipiell beschreibt eine *Simulation* die Nachbildung eines Systems mit seinen dynamischen Prozessen in einem experimentierbaren Modell, um auf die Wirklichkeit übertragbare Erkenntnisse zu erhalten (VDI

3633:05-2018, S. 28). Für PPS gilt die *diskrete ereignisgesteuerte Simulation* (DES) als wichtigste Methode (Rauh 1998, S. 5), bei der Zustandsänderungen über Ereignisse ausgelöst werden (Law 2015, S. 9ff.). Mögliche Ereignisse können bspw. der Beginn oder das Ende einer Bearbeitung sowie die Ein- oder Ausfahrt eines Objektes aus einer Förderstrecke sein (Gutenschwager et al. 2017, S. 54). Als *Materialfluss* wird in diesem Zusammenhang allgemein die Verkettung aller Vorgänge beim Gewinnen, Verteilen, Be- und Verarbeiten von Gütern und Arbeitsgegenständen in einem Produktionssystem bezeichnet (VDI 2689:02-2019, S. 2). In Zeiten von wachsenden Anforderungen an die Flexibilität und Qualität der PPS eröffnen MFS als Entscheidungshilfsmittel so eine Abkehr von der statischen Durchführung produktionstechnischer Planungsaufgaben (Gierth et al. 2006, S. 646f.). Dadurch können kapazitive und zeitliche Zusammenhänge und deren stochastische Störgrößen in verschiedenen Planungsalternativen zur Beurteilung realitätsgetreu und ereignisdiskret dargestellt werden, ohne Störungen im laufenden Betrieb zu verursachen (Rabe 2000, S. 21ff.). Güte und Aufwand einer Simulation hängen maßgeblich vom Grad der *Abstraktion* ab, also der für bestimmte Aspekte der Problemlösung wichtigen Details des Problems (VDI 3633:05-2018, S. 3). Anders als *mathematische Modelle* besitzen MFS den Vorteil, dass sie die systemspezifischen Abhängigkeiten und dynamischen Wechselwirkungen im zeitlichen Verlauf und unter Berücksichtigung stochastischer Aspekte bei der Abbildung des Systemverhaltens berücksichtigen können (März et al. 2011, S. 41f.; Wenzel et al. 2008, S. 10). In später ausgeführten Arbeiten bieten ereignisdiskrete MFS die Grundlage für den Trial-and-Error-basierten Lernansatz der maschinellen Lernverfahren (ML). Ein agentenbasiertes System gewinnt die für die Optimierung notwendige Erfahrung durch vielfache Iterationen der Planungsszenarien, die in einer MFS (anders als in mathematischen Modellen) durch die Einbringung variabilitätsbedingter Änderungseinflüsse deutlich realitätsnäher und damit auch robuster erstellt werden können (Capocchi et al. 2022, S. 3ff.). Ebenso wie *mathematischen Optimierungen*, eignen sich Werkzeuge der *Business Intelligence* lediglich zur Bestimmung optimaler Lösungen bei äußerst simplen Problemen (Waschneck et al. 2018, S. 1264), da sie die eine vollständige Beschreibung aller Wirkzusammenhänge erfordern (Walter et al. 2015, S. 243ff.). Dieses Vorhaben konzentriert sich daher ausschließlich auf die simulationsbasierten Ansätze mittels DES.

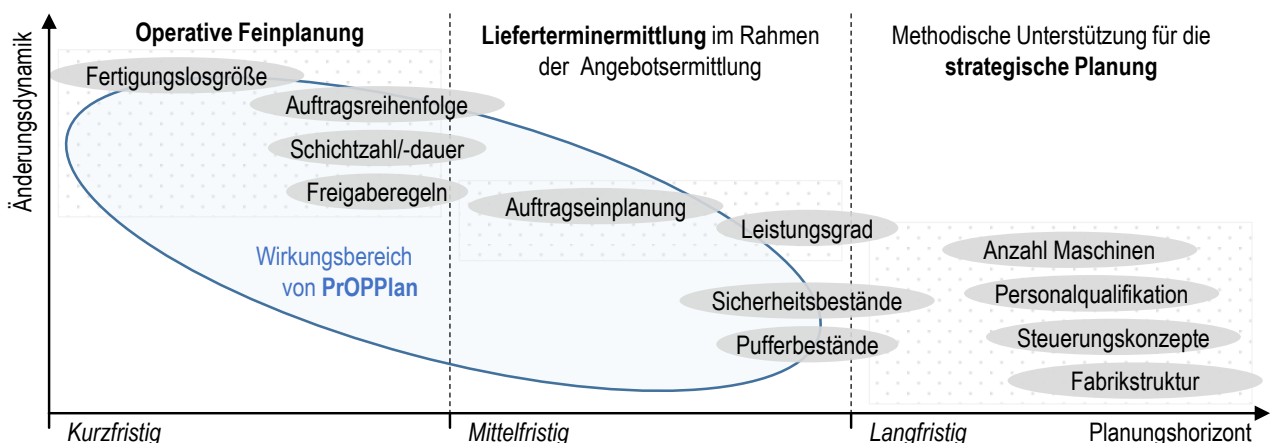


Abbildung 2: Darstellung von Änderungsdynamik und Planungshorizont von produktionstechnischen Aufgaben der simulationsunterstützten Produktionsplanung (Gierth et al. 2006, S. 665) und Verortung des Wirkungsbereichs von PrOPPlan

In Abbildung 2 ist ein Überblick von Planungstätigkeiten der PPP dargestellt, die durch MFS unterstützt werden und die sich nicht nur anhand ihres Planungshorizonts, sondern auch durch die Änderungsdynamiken der Planungsaufgaben unterscheiden (Gierth et al. 2006, S. 664ff.). Die im Vorhaben anvisierten Planungsaufgaben zeichnen sich insb. durch eine hohe Änderungsdynamik und einen kurzen (-bzw. mittelfristigen) Planungshorizont aus. Um Missverständnisse zu verhindern, sei darauf hingewiesen dieser Horizont nicht

den drei zeitlichen Ebenen der PPP (*strategisch, taktisch, operativ*) entspricht. Stattdessen entspricht der Planungshorizont der im Planungsvorgang verfügbaren Zeit zum Umgang mit aufgetretenen Änderungseinflüssen. Durch das Vorhaben werden insb. Produktionssysteme mit hoher Änderungsdynamik unterstützt. Der Einsatz von MFS steht im Zentrum dieses Vorhabens, da sie die dynamischen Abläufe eines Materialflusssystems nachbilden und als Trainings- bzw. Testumgebung für eine intelligente Planungsempfehlung dienen können. ProOPPlan befasst sich daher initial auch mit der Befähigung von Unternehmen, und insb. KMU, zur Nutzung von MFS. Dafür wird auf vorangegangenen Forschungsprojekten aufgebaut, wie bspw. den IGF-Vorhaben KoDeMat (17391 N: *Befähigung von KMU zur kollaborativen Planung und Entwicklung heterogener, dezentral gesteuerter Materialflusssysteme*) oder ProLoPoly (17937 N: *Kybernetisches Simulationsspiel zur Vermittlung abteilungsübergreifender logistischer Zusammenhänge in KMU*), und eine Hilfestellung mit Handlungsempfehlungen entwickelt. Dadurch sollen KMU befähigt werden, MFS zu erstellen und diese mit Hilfe der Methodik zur Erstellung von Planungsempfehlungen wertschöpfend einzusetzen.

2.3 Anwendung von Reinforcement Learning zur selbstlernenden Simulationsoptimierung

Aufgrund der Erfolgsgeschichten populärwissenschaftlicher Anwendungen der *Künstlichen Intelligenz* (KI) erfährt KI als Forschungsgebiet ein seit Jahrzehnten anhaltendes Interesse im gesellschaftlichen Bewusstsein (Bruhn et al. 2021, S. 7). Bspw. gelang es bereits 2013 *DeepMind Technologies* ein selbstlernendes System zu entwickeln, das menschliche Experten in drei von sieben Spielen auf der Arkade-Spielekonsole Atari 2600 schlagen konnte (Mnih et al. 2013, S. 8). Internationale Aufmerksamkeit erregte 2018 der *AlphaZero* Algorithmus, durch den es gelang einen menschlichen Meisterspieler im strategischen Brettspiel Go zu besiegen (Silver et al. 2018, S. 1140ff.). Nur ein Jahr später schlug ein KI-basierter Agent 99,8% alle menschlichen Spieler im Echtzeit-Strategiespiel StarCraft II und gewann auch zuverlässig gegen Spieler vom Rang eines Großmeisters (Vinyals et al. 2019, S. 350ff.). Trotz unterschiedlicher Dimensionen liegt diesen Erfolgen aus einer technischen Perspektive der gleiche Ansatz zugrunde: das *Reinforcement Learning* (RL, engl. für ‚verstärkendes Lernen‘). RL gilt als Teildisziplin des Maschinellen Lernens (Heidrich-Meisner et al. 2007, S. 1f.) und ist gleichzeitig ein Sammelbegriff für eine Reihe von Methoden, bei denen ein Agenten selbstständig eine Strategie erlernt, um eine Belohnung zu maximieren (François-Lavet et al. 2018, S. 224; Sutton et al. 2018, S. 11).

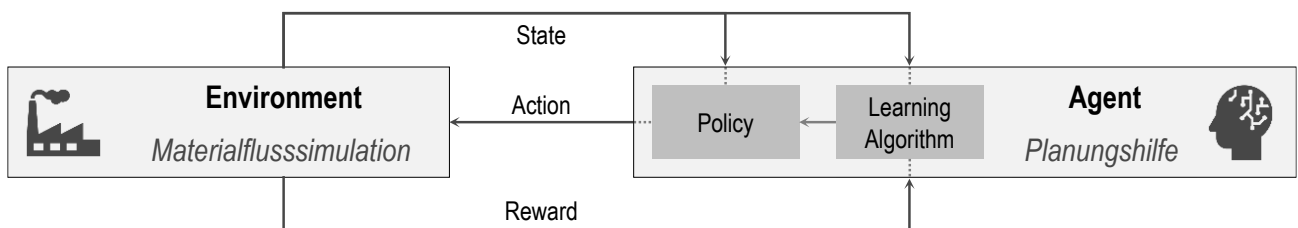


Abbildung 3: Schematisches Wirkprinzip beim Reinforcement Learning (Sutton et al. 2018, S. 38; West et al. 2021, S. 348)

Abbildung 3 veranschaulicht das Wirkprinzip und verortet die wichtigsten Elemente eines RL-Systems im Kontext des vorgesehenen Anwendungsfelds. In ProOPPlan werden MFS als *Environment* (engl. für ‚Umgebung‘) genutzt. Beim Agenten handelt es sich dann um die im Vorhaben zu entwickelnde Planungshilfe für Empfehlungen zur operativen PPP, die wiederum aus zwei Bestandteilen, einer *Policy* (engl. für ‚Richtlinie‘) und dem *Learning Algorithm* (engl. für ‚lernende Handlungsvorschrift‘) besteht. Durch das Environment erhält der Agenten einen als *State* (engl. für ‚Zustand‘) bezeichneten Vektor, der den Ist-Zustand von (allen) verfügbaren Parameter der Simulation enthält. Anhand dieses Zustands muss der Agenten dann eine *Action* (engl. für ‚Handlung‘) aus einer vorgegebenen Menge von Anpassungsmaßnahmen auswählen. Um langfristig gute

Entscheidungen zu verstärken, wird dem Agenten neben dem *State* außerdem ein als *Reward* (engl. für ‚Belohnung‘) bezeichneter Punktwert übermittelt. Dieser Wert wird mithilfe einer Zielfunktion bestimmt und ermöglicht dem lernenden Algorithmus eine Anpassung der Policy (Sutton et al. 2018, S. 7ff.; West et al. 2021, S. 347f.). Auf diese Weise ermöglichen es RL-Anwendungen eine eigenständige und ressourcensparende Lösung für eine kombinatorischen Optimierung zu finden, also aus einer sehr großen Menge diskreter Planungsmöglichkeiten eine optimale (Teil-)Lösung zu konstruieren (Cook et al. 1998, S. 9ff.). RL-basierte Ansätze haben sich in der Vergangenheit außerdem als deutlich weniger aufwändig erwiesen, als detaillierte mathematische Modelloptimierungen. Trotz des populärwissenschaftlichen Interesses und der enormen Potentiale wird im wissenschaftlichen Diskurs ein Mangel ernstzunehmender Anwendungsfälle im produzierende Gewerbe beklagt (Waschneck et al. 2018, S. 1264).

Bereits vor über 25 Jahren beschrieben Zhang et al. (1995, S. 1114ff.) einen Ansatz zur Festlegung von Produktreihenfolgen auf Basis eines neuronalen Netz mit dem sowohl die systemische Ausbringung gesteigert als auch manuelle Planungsaufwände reduziert werden konnten. In den darauffolgenden Jahren haben richtungsweisende Arbeiten, wie bspw. Riedmiller et al. (1999, S. 764ff.) oder Gabel et al. (2007, S. 14ff.), die planungstechnische Überlegenheit von (Deep) RL-Systemen in der PPP aufgezeigt, konnten sich jedoch u.a. aufgrund eines Mangels an verfügbarer Rechenleistung nicht durchsetzen (Panzer et al. 2022, S. 4322). Der einleitend beschriebene Siegeszug von KI-Anwendungen verdeutlicht, dass moderne Industriecomputer längst die Anforderungen der beschriebenen RL-Systeme erfüllen können. Damit einher geht eine starke Zunahme der Forschungsanstrengungen, die in der Anwendung von RL-Systemen in Produktionssystemen unternommen werden: In einer umfassenden Literaturrecherche zeigen Panzer et al. (2022, S. 4320ff.) auf, dass die Anzahl an allgemeinen Publikationen zu RL-Anwendungen in der Produktion jährlich zunimmt und die drei Themenfelder Prozessüberwachung (29 Beiträge), Produktionsprogrammerstellung (22 Beiträge) und Montagedurchführung (18 Beiträge) am präsentesten sind. Damit befindet sich auch die im Vorhaben geplante Methodik unter den derzeit intensiv beforschten Themenfeldern. Die operative PPP zeichnet sich tendenziell durch hohe Unsicherheiten im Hinblick auf kundenindividuelle Produkte, Anlagenausfälle oder sonstige Änderungseinflüsse aus. Um diese Komplexität zu bewältigen und menschliche Entscheidungen zu unterstützen, schlagen Lin et al. (2019, S. 4276ff.) einen Multi-Klassen-DQN-Ansatz vor, also ein auf dem Q-Learning basierendes Modell mit tiefer Netzwerkstruktur, das zur Reihenfolgenplanung in der Halbleiterfertigung eingesetzt wurde und eine bessere Planungsleistung durch reduzierte Auftragsabwicklungszeiten bei gleichbleibenden Durchlaufzeiten erreichte. Um die vergleichsweise hohen Einrichtungs- und Berechnungsaufwände konventioneller Ansätze zur Reihenfolgenplanung zu reduzieren, haben Liu et al. (2020, 71752ff.) und Baer et al. (2019, S. 22ff.) ebenfalls selbstlernende, Multi-Agenten-Systeme entwickelt und konnten lokale sowie globale Produktionsziele erreichen und schnellere Planungsanpassungen erzielen bei einer gleichzeitigen Reduktion der Umplankosten. Ein Jahr darauf demonstrierten Baer et al. (2020, S. 2ff.) die Fähigkeit eines Agenten-Systems, sich an neue Szenarien anzupassen mit guter Skalierbarkeit der Lösung (das Training von 700 Modellen dauerte nur etwa doppelt so lange wie das Training eines einzelnen Modells). Trotz dieser vielversprechenden Arbeiten bilden diese bisher lediglich Insellösungen und besitzen noch weiterführende Handlungsbedarfe. Wenzel et al. (2019, S. 429ff.) zeigen in einer Literaturrecherche Anwendungsbeispiele maschineller Lernverfahren im Supply Chain Management, die nur wenige Anwendungsfälle von Reinforcement Learning beinhaltet. Die Literaturrecherche von Cadavid et al. (2019, S. 385ff.) zum Einsatz von maschinellen Lernansätzen in der PPS bestätigt diese Beobachtung ebenfalls. Eine Zusammenfassung von Kuhnle (2020, S. 53) zeigt darüber hinaus, dass Anwendungsfälle vor allem auf individuelle Planungsaufgaben ausgerichtet sind und der Prämisse kohärenter Planung nicht gerecht werden. Durch die

Entwicklung der Methodik zur selbstlernenden Materialflusssimulation konzentriert sich ProPPlan auf diese Forschungslücke.

3 Durchgeführte Arbeiten und Ergebnisse

Im Projektverlauf wurde eine Methodik zur selbstoptimierenden Produktionsprogrammplanung entwickelt, die Materialflusssimulationen mit Verfahren des Reinforcement Learning kombiniert. Hierzu wurden die relevanten Einflussgrößen operativer Planungsentscheidungen in Produktionssystemen systematisch erfasst, strukturiert und in eine simulationsbasierte Entscheidungsumgebung überführt. Auf dieser Grundlage konnten Planungsaufgaben der operativen Produktionsprogrammplanung – insbesondere Reihenfolgenplanung, Losgrößenbildung, Feinterminierung sowie Auftragsfreigabe – innerhalb eines gemeinsamen Entscheidungsmodells betrachtet werden.

Die entwickelte Methodik ermöglicht damit die Generierung datenbasierter Planungsempfehlungen für komplexe Produktionssysteme und unterstützt eine ganzheitliche Betrachtung konkurrierender Zielgrößen der Produktionsplanung. Durch die Integration selbstlernender Verfahren in eine simulationsgestützte Planungsumgebung wurde ein praxisnahes Instrument geschaffen, das insbesondere kleinen und mittleren Unternehmen eine datengetriebene Entscheidungsunterstützung in der operativen Produktionsplanung ermöglicht. Zur Erreichung der Projektziele wies das Arbeitsdiagramm des Forschungsvorhabens gemäß dem Forschungsantrag folgende Inhalte und Zeiträume aus.

Arbeitspakete (AP) und Teilarbeitspakete (TAP)	Projektlaufzeit: 30 Monate (durch zus. 6 Monate Programmieraufwand in der WDV)																		Aufwand (RIF)						
	2023						2024						2025						[Personenmonate]						
	07	08	09	10	11	12	01	02	03	04	05	06	07	08	09	10	11	12	01	02	03	04	05	06	WMA
1 Ermittlung der PA-Anforderungen und -Bedarfe für Simulationen in Materialflusssystemen	◆																							4,00	0,50
1.1 Definition von planungstechnisch relevanten Anforderungen an die Produktionsprogrammplanung																								1,00	0,25
1.2 Definition systemischer Anforderungen und Befähigung zur Verwendung von Materialflusssimulationen																								3,00	0,25
2 Ableitung der relevanten Produktionsparameter für selbstlernende Simulationsumgebungen										◆														6,00	1,50
2.1 Ermittlung geeigneter Anpassungsmaßnahmen zur Beeinflussung der Materialflusssysteme (Action)																								1,00	0,25
2.2 Ermittlung notwendiger Zustandsgrößen zur Rückmeldung von Planungsanpassungen (State)																								2,00	0,50
2.3 Entwicklung einer parametrierbaren Belohnungsfunktion zur zielgerichteten Optimierung (Reward)																								3,00	0,75
3 Entwicklung der Methodik zur selbstoptimierenden Simulation mittels Reinforcement Learning																								10,00	2,00
3.1 Pilotierende Entwicklung und Anwendung der Methodik für einen Anwendungsfall aus dem PA																								4,00	0,75
3.2 Reproduktion der Methodik für weitere Anwendungsfälle und Generalisierung der Vorgehensweise																								4,00	0,75
3.3 Entwicklung von Maßnahmen zur Gewährleistung der Nachvollziehbarkeit von Planungsempfehlungen																								2,00	0,50
4 Anwendung und Optimierung der entwickelten Methodik in realen Materialflusssystemen																								4,00	1,00
4.1 Vorbereitung und Begleitung während der Einführung der entwickelten Methodik bei den Anwendern																								2,00	0,50
4.2 Untersuchung und abschließende Bewertung der Einführung hinsichtlich der initialen Zielsetzung																								2,00	0,50
5 Überführung der Methodik in ein IT-Werkzeug und vorhabenbegleitende Optimierung																								6,00	1,00
6 Projektmanagement, Öffentlichkeitsarbeit, Dokumentation und Ergebnistransfer	Arbeitspaket 6 wird anteilig in den Arbeitspaketen 1 bis 5 durchgeführt.																		---	---					
Σ Sitzungen des projektbegleitenden Ausschusses und Meilensteine (MS)	MS1: Workshop zur Anforderungserfassung			MS2: Erste Umsetzung und Parameterfestlegung			MS3: Beginn der IT-Werkzeug-Entwicklung			MS4: Einführung des Werkzeugs; Abschluss			30,00	6,00											

Abbildung 4: Projektplan gemäß Antrag

Auf Grund einer erschwerten Personalgewinnung im Jahr 2023, die insbesondere in der frühen Projektphase zu einer reduzierten Bearbeitungskapazität führte, konnten die Inhalte der Arbeitspakete AP 1 und AP 2 erst verzögert abgeschlossen werden. Um den entstandenen Rückstand zu kompensieren, wurde eine kostenneutrale Laufzeitverlängerung um sechs Monate beantragt und bewilligt.

Nach erfolgreichem Abschluss der Personalakquise konnte die Forschungsarbeit im Januar 2024 planmäßig aufgenommen werden. Um die wissenschaftlichen Ziele des Vorhabens in der vorgesehenen Tiefe zu erreichen, wurde im Januar 2025 eine ausgabenneutrale Laufzeitverlängerung um sechs Monate bis zum 31.12.2025 beantragt. Mit dieser Maßnahme bleibt die vollständige Erreichung der Projektziele unter Berücksichtigung der ursprünglichen inhaltlichen Zielsetzungen gewährleistet.

Die Bearbeitung der Arbeitspakete erfolgte in enger Abstimmung mit den im projektbegleitenden Ausschuss engagierten Industriepartnern. Die Forschungsaktivitäten wurden kontinuierlich an den praktischen Anforderungen der beteiligten Unternehmen ausgerichtet. Insbesondere bei der Entwicklung der Methodik sowie bei der prototypischen Umsetzung des IT-Werkzeugs erwies sich ein iteratives Vorgehen unter Einbindung realer Anwendungsfälle als zielführend, die Bearbeitung des AP 5 wurde daher vorgezogen, um die Interaktion mit den PA-Mitgliedern effektiver zu gestalten und die Methodik iterativ zu entwickeln.

Die Zielerreichung des Forschungsvorhabens war dadurch jedoch nicht gefährdet. Im Gegenteil konnten durch die intensive Einbindung der Industriepartner sowie durch die iterative Weiterentwicklung der Methodik zusätzliche Erkenntnisse gewonnen und die Praxistauglichkeit der Ergebnisse weiter verbessert werden.

Mit Abschluss des Forschungsvorhabens liegen die angestrebten Ergebnisse aller Arbeitspakete vor. Die entwickelte Methodik zur selbstoptimierenden Produktionsprogrammplanung sowie das prototypisch umgesetzte IT-Werkzeug ermöglichen eine simulationsgestützte Generierung datenbasierter Planungsempfehlungen für die operative Produktionsplanung. Damit steht ein praxisnaher Ansatz zur Verfügung, der Unternehmen bei der Bewältigung komplexer Planungsaufgaben unterstützt und die Nutzung moderner datengetriebener Methoden in der Produktionsplanung fördert.

3.1 Arbeitspaket 1: Ermittlung der PA-Anforderungen und -Bedarfe für Simulationen in Materialflusssystemen

Ziel des Arbeitspakets 1 war die systematische Erhebung und Strukturierung der Anforderungen und Bedarfe des PA im Hinblick auf den Einsatz selbstoptimierender Materialflusssimulationen zur operativen Produktionsprogrammplanung (PPP). Im Fokus standen sowohl planungstechnisch relevante Anforderungen an die operative PPP als auch systemische Anforderungen zur Befähigung von KMU zur Einführung von Simulation und RL. Das Arbeitspaket bildet damit die konzeptionelle Grundlage für die nachfolgenden Entwicklungsarbeiten (insbesondere Definition von State, Action und Reward sowie Ableitung der Produktionsparameter).

Durchgeführte Arbeiten

Im Rahmen des Arbeitspakets wurden zunächst die Anforderungen der PA Mitglieder systematisch erhoben. Hierzu wurde zu Projektbeginn ein Kick-off-Workshop durchgeführt, in dem die Zielsetzung des Vorhabens, zentrale Begrifflichkeiten sowie potenzielle Einsatzfelder einer selbstoptimierenden Produktionsprogrammplanung präzisiert wurden. Aufbauend darauf erfolgte eine strukturierte Diskussion der Kernaufgaben der operativen Produktionsprogrammplanung – insbesondere Feinterminierung, Losgrößenrechnung, Reihenfolgeplanung und Auftragsfreigabe – sowie ihrer praktischen Herausforderungen. Die hohe kombinatorische Komplexität dieser Planungsprobleme wurde exemplarisch verdeutlicht, um den Bedarf an lernfähigen Optimierungsansätzen nachvollziehbar darzustellen.

Im Anschluss an den Workshop wurden in weiteren Gesprächen und Vor-Ort-Terminen mit den PA-Partnern die spezifischen Erwartungen an die geplante Methodik vertieft. Parallel hierzu erfolgte eine systematische Auswertung einschlägiger wissenschaftlicher und praxisorientierter Literatur zur PPS, MFS sowie zu Einführungsstrategien für RL im industriellen Umfeld. Ziel war es, die praxisbezogenen Anforderungen mit dem Stand der Forschung abzugleichen und strukturiert aufzubereiten.

Auf dieser Grundlage wurde ein strukturierter Anforderungskatalog erarbeitet, der zunächst planungstechnisch relevante Anforderungen an eine RL-basierte operative Produktionsprogrammplanung umfasst. Hierzu

zählen insbesondere Anforderungen an die Qualität der Planungsergebnisse im Vergleich zu bestehenden PPS-, MES- oder APS-Systemen, die Möglichkeit der Optimierung nach unterschiedlichen – teils konkurrierenden – Zielgrößen (z. B. Termintreue, Durchlaufzeit, Auslastung, Robustheit), die Berücksichtigung branchenspezifischer technischer Restriktionen sowie die Integration von Rüstzeit- und Rüstkostenaspekten in die Reihenfolgeplanung. Darüber hinaus wurden Anforderungen an eine benutzerfreundliche Schnittstelle, flexible Zielgrößenkonfiguration, einfache Datenintegration aus bestehenden IT-Systemen sowie praktikable Rechenzeiten (z. B. Nachtläufe oder Kurzoptimierungen) erfasst.

Ergänzend wurde der Anforderungskatalog um simulationstechnische und systemische Aspekte erweitert. Hierbei wurden Anforderungen an die Definition von Systemgrenzen, die Qualität und Verfügbarkeit von Produktions- und Auftragsdaten, die Validierung von Simulationsmodellen sowie an die organisatorischen Rahmenbedingungen einer Einführung berücksichtigt. Zudem wurde eine erste Anwendergruppierung vorgenommen, um unterschiedliche Problemcharakteristika – beispielsweise variantenreiche Fertigung, hohe Automatisierung oder begrenzte IT-Ressourcen – systematisch zu erfassen. Aufbauend auf den Erkenntnissen aus Praxis und Literatur wurde schließlich ein KMU-zentrierter Leitfaden zur strukturierten Einführung von Simulation und RL entwickelt, der ein phasenbasiertes Vorgehensmodell von der Vorbereitung über die Umsetzung bis zur Integration in den Echtbetrieb beschreibt.

Erzielte Ergebnisse

Als wesentliches Ergebnis von Arbeitspaket 1 liegt ein strukturierter und validierter Anforderungskatalog vor, der sowohl planungstechnische als auch systemische Anforderungen an eine selbstoptimierende Materialflusssimulation abbildet. Dieser Katalog bildet die zentrale Referenz für die weitere methodische Ausgestaltung des Vorhabens und dient insbesondere als Grundlage für die spätere Definition der Zielfunktion (Reward), der relevanten Zustandsgrößen (State) sowie der zulässigen Eingriffs- und Steuerungsoptionen (Action).

Ein weiteres zentrales Ergebnis ist die Erkenntnis, dass in der operativen Produktionsprogrammplanung keine singuläre Optimierungsgröße existiert, sondern mehrere, teils konkurrierende Zielgrößen simultan berücksichtigt werden müssen. Daraus ergibt sich die Anforderung einer konfigurierbaren, parametrierbaren Zielfunktionsgestaltung, um unterschiedliche betriebliche Prioritäten abbilden zu können. Ebenso wurde herausgearbeitet, dass technische Restriktionen und branchenspezifische Besonderheiten nicht als Randbedingungen zweiter Ordnung, sondern als integraler Bestandteil der Modellierung zu berücksichtigen sind.

Darüber hinaus wurde mit dem entwickelten Leitfaden ein strukturierter Einführungsrahmen für KMU geschaffen, der technische, organisatorische und wirtschaftliche Aspekte integriert. Damit wurde sichergestellt, dass die im Projekt entwickelten Methoden nicht ausschließlich algorithmisch fundiert, sondern auch praxisgerecht implementierbar sind. Dieser ist auf der Homepage des RIF verfügbar. Arbeitspaket 1 schafft somit die konzeptionelle und anwendungsbezogene Grundlage für die nachfolgenden Arbeitspakete und stellt sicher, dass die weitere Entwicklung eng an den realen Bedarfen der Anwenderunternehmen ausgerichtet ist.

3.2 Arbeitspaket 2: Ableitung der relevanten Produktionsparameter für selbstlernende Simulationsumgebungen

Durchgeführte Arbeiten

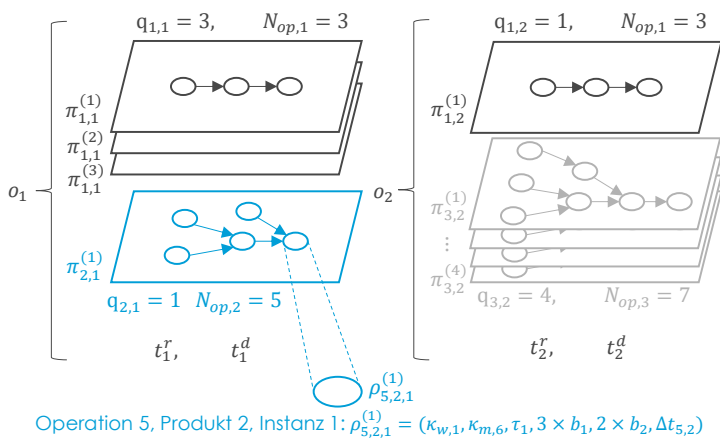
Die relevanten Produktionsparameter für selbstlernende Simulationsumgebungen wurden auf Grundlage wissenschaftlicher Arbeiten sowie durch Abstimmung mit den beteiligten Unternehmen erarbeitet. Maßgeblich dabei war die Dissertation von Rinciog (2023), welche zeigt, dass RL als methodischer Kompromiss zwischen Prioritätsregeln und exakten Optimierungsverfahren gilt. Für eine fundierte Parametersauswahl der selbstlernenden Simulationsmodelle wurden mittels umfassender Literaturlauswertung sowie praxisnahen Diskussion mit den Unternehmen die relevanten Einflussgrößen identifiziert.

Erzielte Ergebnisse

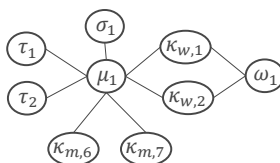
Als zentrale Entscheidungsdimensionen (*Actions*) der operativen PPP wurden Routing (Operationszuteilung) und Sequencing (Operationsauswahl) identifiziert, wobei auch der bewusste Verzicht auf eine Entscheidung als strategische Option berücksichtigt wurde. Es wird zwischen direkten Maßnahmen (detaillierte Steuerung einzelner Aufträge und Ressourcen) mit potenziell global optimalen, aber datenintensiven Lösungen und einfacher umsetzbaren Maßnahmen unterschieden (Anpassung von Prioritätsregeln wie FIFO oder KOZ). Die Wirksamkeit beider Ansätze hängt von den Zustandsgrößen (*States*) ab, die zwischen vollständigem Systemzustand (*environment-state*) und begrenzter Beobachtung des Agenten (*agent-state*) unterschieden werden. Die daraus resultierende partielle Beobachtbarkeit stellt insbesondere für KMU eine Herausforderung dar, weshalb ereignisdiskrete Simulationen zur Ergänzung der notwendigen Informationen eingesetzt werden.

Anfang Planungsperiode: t_0
 Ende Planungsperiode: t_e
 Aktueller Zeitpunkt: $t_0 \leq t \leq t_e$

Kundenaufträge (Produktvorranggraphen):



Produktionsressourcen:



σ : Arbeitsstation
 μ : Maschine
 ω : Werker / Mitarbeiter
 τ : Werkzeug / Hilfsmittel
 κ_w : Mitarbeiterfähigkeit
 κ_m : Maschinenfähigkeit

Beobachtungsraum (Matrixform):

	Arbeitsstationen									
	σ_1	σ_2	$\sigma_{N_{ws}}$
$\rho_{x,y,z}^{(c)}$	1	0	0
\vdots	0	0	0
\vdots	0	1	0
\vdots										
τ_1	0.333	0.000	0.000
\vdots	0.000	0.667	0.000
\vdots	0.000	1.000	0.000
\vdots										
τ_{N_t}										
ω_1	1.000	0.000	0.000
\vdots	0.000	0.000	0.000
\vdots	0.000	0.750	0.000
\vdots										
ω_{N_w}										

Abbildung 5: Mathematische Formalisierung der Produktionssysteme erforderlich für Reinforcement Learning

Die nach Rinciog (2023) definierten Zustandsgrößen wurden um planungstechnisch relevante Variablen erweitert, um die Komplexität des Zustandsraums im Programmcode abbilden zu können. Abbildung 5 demonstriert das entwickelte mathematische Modell, das in AP 5 als Grundlage für die Simulationslogik diente.

Anhand der Abbildung 5 werden im Folgenden die für die operative PPP (Feinterminierung, Losgrößenrechnung, Reihenfolgeplanung und Auftragsfreigabe) erforderlichen Zustandsgrößen erklärt.

Zur termingerechten Fertigung müssen Produktionssysteme alle Kundenaufträge, bestehend aus Produkten und mehreren möglicherweise in Präzedenzbeziehung stehenden Operationen, bewältigen. Bei ausreichenden Ressourcen können Operationen parallelisiert werden, insbesondere bei vorhandenen Baugruppenstrukturen der Produkte. Die Zusammensetzung der Kundenaufträge ist in Abbildung 5 mit dem Symbol o gekennzeichnet.

Operationen p erfordern spezifische Ressourcen wie Mitarbeiter, Maschinen, Werkzeuge und Materialien sowie eine definierte Bearbeitungszeit Δt (vgl. Kluge 2011). Die Zuordnung der Operationen zu den richtigen Zeitpunkten bildet den Kern der Produktionssteuerung und ermöglicht eine simultane Abbildung von Feinterminierung, Reihenfolgeplanung, Losgrößenrechnung und Auftragsfreigabe.

Die Produktionsressourcen und ihre Wechselwirkungen sind im unteren linken Bildbereich der Abbildung 5 skizziert. Ein Kernelement von Produktionssystemen sind Arbeitsstationen (AS), die sich aus verschiedenen Entitäten zusammensetzen (vgl. Abbildung 6).

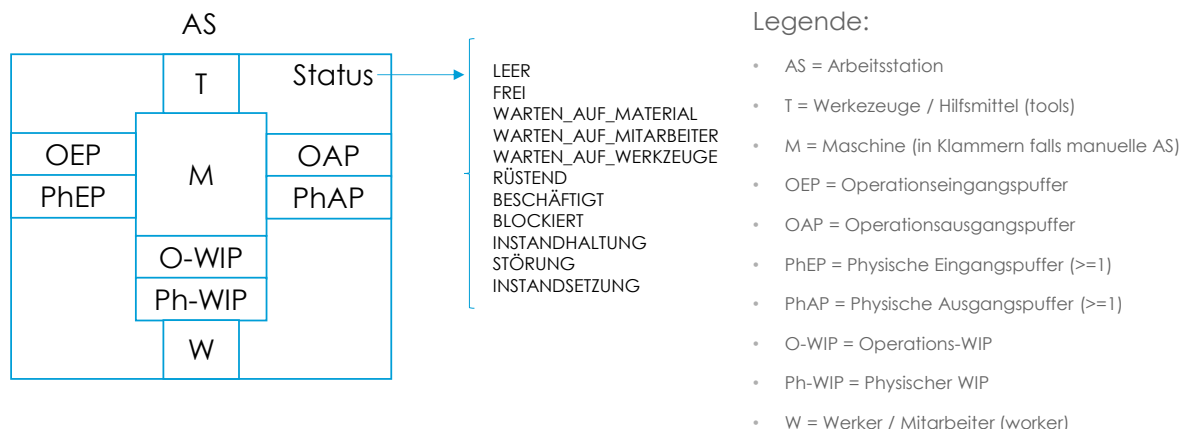


Abbildung 6: Arbeitsstation als zentrales Element von Produktionssystemen

Das entwickelte Modell bildet unterschiedliche Fertigungsformen ab und steuert den Materialfluss über die Zuordnung von Fähigkeitsanforderungen zu Ressourcen. Der RL-Zustandsraum ist als Matrix organisiert, was die Berücksichtigung von Kapazitätsrestriktionen und die Integration logistischer Zielgrößen erlaubt. Zusätzlich werden verdichtete Kennzahlen aus den Rohdaten abgeleitet, die Auftrags-, Ressourcen- und Zielgrößenbezogene Parameter umfassen. Die finale Wirksamkeit der Zustandsvariablen erfordert eine Validierung in realen Anwendungsszenarien, weshalb eine konfigurierbare Modellstruktur priorisiert wird.

Die Belohnungsfunktion (*Reward*) informiert den RL-Agenten über die Zielerreichung eines Zustands und steuert damit das Lernverhalten. Die Definition der Optimierungsziele bildet die Grundlage für ihre Ausgestaltung. Diese Ziele können sich jedoch für jedes Unternehmen unterscheiden. In diesem Fall wurden sie in Abstimmung mit dem PA-Mitglied BMW AG unter Rückgriff auf etablierte Ansätze konkretisiert (vgl. Ostermeier et al. 2024). Typische Zielgrößen sind die Durchlaufzeit, Liefertermintreue, Bestände und Auslastung

Aufgabe der Produktionsprogrammplanung ist es, das Produktionsprogramm des Unternehmens zu erstellen. Dieses enthält für jedes Erzeugnis und für jede Planungsperiode die zu erstellende Menge. Es legt also den Netto-Primärbedarf eines Unternehmens fest (Lödding 2016, S. 108).

3.3 Arbeitspaket 3: Entwicklung der Methodik zur selbstoptimierenden Simulation mittels RL

Durchgeführte Arbeiten

Im betrachteten Arbeitspaket wurde die in den vorangegangenen Arbeitspaketen konzipierte Methodik zur selbstoptimierenden Produktionsprogrammplanung zunächst pilotierend auf einen realitätsnahen industriellen Anwendungsfall angewendet und anschließend auf einen zweiten Use Case übertragen. Damit folgte die Bearbeitung der im Forschungsantrag vorgesehenen Logik, wonach die Methodik zunächst an einem geeigneten Pilotfall entwickelt und anschließend an weiteren Anwendungsszenarien reproduziert und generalisiert werden sollte.

Den Ausgangspunkt bildete ein erster industrieller Use Case aus der Gehäusefertigung eines Pumpenherstellers. Auf dessen Basis wurde für diesen Anwendungsfall ein anpassbares, mehrstufiges Fertigungsszenario aufgesetzt, das die wesentlichen Bearbeitungsschritte vom Wareneingang über Lager, mechanische Bearbeitung und Oberflächenbehandlung bis hin zur Montage und zum Warenausgang umfasste. Innerhalb der einzelnen Bearbeitungsstufen wurden mehrere parallele Maschinen berücksichtigt, wobei gleichartige Operationen in Abhängigkeit vom jeweiligen Produkt unterschiedlich lange Bearbeitungszeiten aufwiesen. Darüber hinaus wurde modelliert, dass nicht jede Station technisch zur Bearbeitung aller Produkte geeignet ist. Zusätzlich wurden produktabhängige Rüstzeiten sowie Konzepte der Mehrmaschinenbedienung in den relevanten Fertigungsbereichen berücksichtigt. Auf diese Weise entstand ein realitätsnahes und zugleich variierbares Planungsszenario, das sich insbesondere zur initialen Erprobung der entwickelten Methodik unter unterschiedlichen Auftragslagen und Rüstsituationen eignete.

Im Zuge dieser Pilotierung wurden die für die Methodik erforderlichen Modellbestandteile iterativ konkretisiert. Dazu gehörten insbesondere die Definition der planungsrelevanten Zustandsgrößen, die Ausgestaltung der zulässigen Entscheidungs- bzw. Anpassungsmaßnahmen innerhalb des Planungssystems sowie die Parametrisierung der Zielfunktion. Parallel dazu wurde geprüft, in welchem Detaillierungsgrad der reale Fertigungsprozess abgebildet werden musste, um sowohl eine hinreichende Realitätsnähe als auch eine praktikable Modellkomplexität sicherzustellen. Die Arbeiten umfassten somit nicht nur die technische Abbildung des Pilotfalls im Simulationsmodell, sondern auch die methodische Ausgestaltung der eigentlichen Entscheidungslogik für die operative Produktionsprogrammplanung.

Aufbauend auf diesen Arbeiten wurde die entwickelte Methodik in einem zweiten Schritt auf einen weiteren industriellen Use Case übertragen. Hierfür wurde ein Anwendungsfall aus der Medizintechnik herangezogen, bei dem ein Hersteller von Medizinprodukten seine Endmontage von einer klassischen Baustellen- bzw. Inselstruktur in eine fluss- bzw. linienorientierte Fertigung überführt hatte. Dieser Transformationsprozess schuf besonders geeignete Rahmenbedingungen für die Reproduktion der Methodik, da im Zuge der Umstellung mehrere operative Planungsaufgaben neu durchdacht werden mussten. Dies betraf insbesondere die Einplanung von Aufträgen in eine getaktete Endmontage, die Zuordnung von Kapazitäten zu einzelnen Stationen, die Stabilisierung des Materialflusses sowie die Frage nach einem geeigneten Umgang mit schwankenden Arbeitsinhalten und variierender Auslastung.

Für diesen zweiten Use Case wurden zunächst die Struktur des neuen Montagesystems, die Taktlogik der Linie, vorgelagerte Vorfertigungsprozesse sowie die maßgeblichen Restriktionen der betrieblichen Planung aufgenommen und in ein geeignetes Simulationsmodell überführt. Anschließend wurde geprüft, inwieweit die im ersten Use Case entwickelte Methodik mit vertretbarem Anpassungsaufwand auf den neuen Anwendungskontext übertragen werden konnte. Ein Schwerpunkt der Arbeiten lag dabei auf der Reproduzierbarkeit des methodischen Vorgehens. Es wurde also bewusst nicht allein ein weiterer Einzelfall modelliert, sondern untersucht, welche Elemente der Methodik unmittelbar übertragbar waren, an welchen Stellen anwendungsspezifische Anpassungen erforderlich wurden und wie robust die Modelllogik gegenüber veränderten Produktionsstrukturen reagierte.

Ergänzend wurden für beide Use Cases vergleichbare Basisszenarien und Referenzannahmen aufgebaut, um die spätere Bewertung der Planungsempfehlungen auf eine konsistente Grundlage zu stellen. Zudem erfolgte eine enge fachliche Abstimmung mit den jeweiligen Praxispartnern, um reale Restriktionen, typische Planungsprobleme und betrieblich relevante Zielgrößen adäquat in den Modellen abzubilden. Insgesamt diente das Arbeitspaket damit der methodischen Konsolidierung des entwickelten Ansatzes: Während der erste Anwendungsfall vorrangig der Pilotierung und dem Aufbau eines belastbaren Simulations- und Entscheidungsrahmens diente, stand beim zweiten Anwendungsfall die Reproduktion, Übertragbarkeit und praxisnahe Validierung der Vorgehensweise im Vordergrund.

Im Rahmen von TAP 3.3 wurden Maßnahmen zur Erhöhung der Transparenz und Nachvollziehbarkeit der erzeugten Planungsempfehlungen entwickelt. Hintergrund war die im Antrag hervorgehobene Notwendigkeit, RL-basierte Modelle trotz ihres Black-Box-Charakters für industrielle Anwender verständlich und akzeptabel zu machen. Hierzu wurden insbesondere Trainingsverläufe visualisiert, RL-Ergebnisse systematisch mit heuristischen Referenzlösungen verglichen und die wesentlichen Versuchsparameter nachvollziehbar dokumentiert. Die dabei gewonnenen Erkenntnisse flossen zudem in die nutzerorientierte Ausgestaltung des IT-Werkzeugs sowie in den projektbegleitend entwickelten Leitfaden ein.

Erzielte Ergebnisse

Im Rahmen des Pilot-Use-Cases konnte die entwickelte Methodik erstmals anhand eines realitätsnahen, zugleich aber kontrolliert variierbaren Fertigungsszenarios angewendet und bewertet werden. Hierzu wurden unterschiedliche Auftrags- und Rüstsituationen in zwei Varianten abgebildet. Während das einfachere Szenario v5 vier Kundenaufträge, 87 zu produzierende Erzeugnisse und 510 einzuplanende Operationen umfasste, wurde mit dem komplexeren Szenario v7 ein deutlich anspruchsvollerer Fall mit sechs Kundenaufträgen, 341 Erzeugnissen und 2.010 Operationen betrachtet. Zusätzlich unterschieden sich die Szenarien durch das Niveau der produktabhängigen Rüstzeiten, das von wenigen Minuten in v5 bis zu deutlich höheren Rüstzeitintervallen in v7 reichte. Damit stand ein geeigneter Pilotfall zur Verfügung, um die entwickelte Methodik unter unterschiedlich komplexen Randbedingungen zu erproben.

Die durchgeführten Experimente zeigten zunächst, dass sich der Anwendungsfall sehr gut für ein schnelles vergleichendes Testen unterschiedlicher Planungslogiken eignete. Auf Basis des Simulationsmodells konnten verschiedene heuristische Regelkombinationen sowie erste RL-Konfigurationen effizient untersucht und gegenübergestellt werden. Im heuristischen Vergleich erwiesen sich insbesondere Kombinationen aus geeigneter Arbeitsplatzzuordnung und einer prioritätsorientierten Reihenfolgebildung als leistungsfähig. Die beste im Pilotfall identifizierte Heuristik erreichte eine durchschnittliche Auftragsdurchlaufzeit von rund 6 h 5 min und stellte damit die relevante Referenz für die nachfolgenden RL-Experimente dar.

Die RL-basierten Versuche, durch geführt auf einem leistungsfähigen, lokalen Server, zeigten, dass der entwickelte Ansatz grundsätzlich lernfähig war und sich gegenüber rein zufälligen Entscheidungen verbessern konnte. Bei einer Konfiguration mit 50 Simulationsläufen wurde für das beste RL-Ergebnis eine durchschnittliche Durchlaufzeit von rund 6 h 27 min erzielt. Dieses Ergebnis lag zwar über typischen Zufallsresultaten, erreichte jedoch nicht die Qualität der besten heuristischen Referenzlösung. Auch bei einer Erhöhung der Trainingsintensität auf 500 Simulationsläufe konnte keine verlässlich bessere Lösung identifiziert werden. Vielmehr zeigte sich eine deutliche Sensitivität gegenüber der gewählten Konfiguration, den Beobachtungs- und Aktionsräumen, der Netzarchitektur sowie den weiteren Hyperparametern des Lernverfahrens. Verlängerte Trainingsdauern führten im Pilotfall ebenfalls nicht zu einer robusten Leistungssteigerung, sondern erhöhten vor allem den rechnerischen Aufwand erheblich.

Als wesentliches Ergebnis des Pilot-Use-Cases ist damit festzuhalten, dass die methodische Kopplung von Materialflusssimulation und Reinforcement Learning erfolgreich umgesetzt und experimentell untersucht werden konnte. Gleichzeitig wurde deutlich, dass RL in dem betrachteten industriellen Szenario zwar in einzelnen Konfigurationen, aber nicht automatisch immer zu besseren Planungsergebnissen führt als gut abgestimmte heuristische Verfahren. Der Pilotfall machte somit zweierlei sichtbar: Zum einen ist der entwickelte methodische Rahmen geeignet, unterschiedliche Planungslogiken systematisch zu vergleichen und RL-basierte Ansätze technisch zu integrieren. Zum anderen sind die Leistungsfähigkeit und Wirtschaftlichkeit des RL-Einsatzes stark von der konkreten Problemstruktur, der Güte der Modellierung sowie dem vertretbaren Trainingsaufwand abhängig. Für das Vorhaben ergab sich daraus die wichtige Erkenntnis, dass heuristische Referenzlösungen im weiteren Projektverlauf nicht nur als Benchmark, sondern auch als praxisrelevanter Vergleichsmaßstab für die Bewertung datengetriebener Ansätze zwingend mitzudenken sind.

Im zweiten Anwendungsfall wurde die entwickelte Methodik auf ein industrielles Szenario aus der Medizintechnik übertragen. Der Use Case betraf den Hersteller von Medizinprodukten.

Der Use Case erwies sich als gut geeigneter Praxisfall, um den entwickelten Simulationskern in einem weiteren industriellen Kontext anzuwenden und konkrete betriebliche Fragestellungen zu analysieren. Hierzu wurde zunächst eine Initialplanung für einen Betrachtungshorizont von einer Arbeitswoche erstellt. Aufbauend darauf wurden alternative Planungsszenarien untersucht, insbesondere hinsichtlich unterschiedlicher Personal-Station-Zuordnungen und reduzierter Personalbelegungen.

Die Analysen zeigten, dass das betrachtete Wochenprogramm auch bei einer Reduktion der Personalbelegung von sieben auf sechs Mitarbeitende grundsätzlich realisierbar blieb. Gleichzeitig führte dieses Szenario jedoch zu einer deutlichen Erhöhung der Durchlaufzeit sowie zu einer Verschiebung von Engpässen innerhalb des Systems. Darüber hinaus stiegen die Umlaufbestände und damit die Kapitalbindung, während sich die Auslastung einzelner Stationen erhöhte. Damit konnten die betrieblichen Auswirkungen alternativer Planungsentscheidungen transparent gemacht und quantitativ bewertet werden.

Für den betrachteten Anwendungsfall zeigte sich zugleich, dass der Einsatz von Reinforcement Learning nicht zielführend war. Aufgrund der vergleichsweise geringen kombinatorischen Komplexität konnten die wesentlichen Erkenntnisse bereits durch die strukturierte Analyse heuristisch erzeugter Alternativszenarien gewonnen werden. Der Use Case machte damit deutlich, dass die entwickelte Methodik nicht nur Potenziale, sondern auch Grenzen des RL-Einsatzes in der operativen Produktionsplanung sichtbar machen kann.

Darüber hinaus wurde im Projekt deutlich, dass der entwickelte Tool-Kern auch für kurzfristige Planungsanpassungen ein relevantes Anwendungspotenzial besitzt, etwa bei Lieferverzögerungen, Personalausfällen,

Qualitätsproblemen oder kurzfristigen Änderungen der Auftragslage. Der Medizintechnik-Use-Case bestätigte damit die praktische Übertragbarkeit des entwickelten Ansatzes und zeigte zusätzliche Einsatzmöglichkeiten für simulationsgestützte Planung in realen betrieblichen Entscheidungssituationen auf.

Aus der pilotierenden Anwendung und der Reproduktion der Methodik in zwei unterschiedlichen industriellen Kontexten konnten über die Einzelfälle hinaus verallgemeinerbare Erkenntnisse für die Einführung simulations- und RL-gestützter Produktionsplanung in KMU abgeleitet werden. Die Arbeiten zeigten, dass sich ein strukturiertes Vorgehen in drei Phasen bewährt: Zunächst sind Problemstellung, Zielgrößen, Systemgrenzen und Datenverfügbarkeit präzise zu definieren. Darauf aufbauend erfolgt die Umsetzung durch Modellierung und Validierung des Simulationsmodells sowie – bei hinreichender Problemkomplexität – die Kopplung mit einem RL-Agenten. In einer dritten Phase steht die betriebliche Integration im Vordergrund, insbesondere die Nutzung der entwickelten Logik für Szenarioanalysen, schrittweise Planungsunterstützung und perspektivisch für die operative Anwendung.

Darüber hinaus wurde deutlich, dass der Nutzen des entwickelten Ansatzes nicht allein im Einsatz von Reinforcement Learning liegt. Vielmehr erwiesen sich ein klar abgegrenzter Pilotprozess, eine belastbare Datenbasis, die enge Einbindung von Prozessexperten sowie die Möglichkeit zur schrittweisen Skalierung als zentrale Erfolgsfaktoren. Die Projektergebnisse stützen damit die im Leitfaden formulierte Empfehlung, Simulation und RL nicht als isolierte Technologien, sondern als methodisch eingebettete Werkzeuge zur datengetriebenen Verbesserung der Produktionsplanung zu verstehen.

Als Ergebnis zeigte sich außerdem, dass Nachvollziehbarkeit im vorliegenden Anwendungskontext vor allem durch eine transparente Darstellung von Trainingsfortschritt, Parameterwahl und Benchmark-Ergebnissen gefördert werden konnte. Die Gegenüberstellung von RL, Heuristiken und Zufall erleichterte die Einordnung der Planungsempfehlungen und machte sowohl Potenziale als auch Grenzen des RL-Einsatzes sichtbar. Damit wurde ein praxisgerechter Beitrag zur Akzeptanz der entwickelten Methodik geleistet, wie im Antrag für TAP 3.3 vorgesehen.

3.4 Arbeitspaket 4: Anwendung und Optimierung der entwickelten Methodik in realen Materialflusssystemen

Durchgeführte Arbeiten

Im Rahmen der anwendungsnahen Weiterentwicklung und Erprobung des entwickelten Tools wurde ein weiterer industrieller Use Case bei einem Hersteller von Großgeräten betrachtet. Ziel dieses Anwendungsfalls war die simulationsgestützte Materialplanung für einen Wochenhorizont. Ausgangspunkt war die betriebliche Situation, dass bislang keine systematische Vorausplanung des Materialbedarfs auf Basis realer Produktionsaufträge erfolgte. Vor diesem Hintergrund wurde das Tool so eingesetzt und weiterentwickelt, dass reale Aufträge einer Kalenderwoche simuliert und die daraus resultierenden Materialbedarfe im Voraus abgeschätzt werden konnten.

Für die Anwendung wurden reale Produktionsaufträge aus dem ERP-System des Unternehmens herangezogen. Hierzu wurde eine Schnittstellenlogik umgesetzt, die nach Auswahl eines Betrachtungszeitraums den Verbindungsaufbau zum ERP-System, die Formatkonvertierung der Auftragsdaten sowie den Import in das PrOPPlan-Datenformat ermöglichte. Die importierten Aufträge wurden anschließend mit einem festgelegten Satz von Simulationseinstellungen verarbeitet. Dabei kamen für die Arbeitsplatzzuordnung, die Reihenfolgebildung sowie für Routing- und Sequenzierungsentscheidungen fest definierte indirekte Regeln zum Einsatz.

Auf diese Weise konnte eine stabile und reproduzierbare Ausgangsbasis für die simulationsgestützte Wochenbetrachtung geschaffen werden.

Begleitend zur Anwendung wurden mehrere technische Erweiterungen am Tool vorgenommen, um dessen Praxistauglichkeit und Auswertbarkeit zu erhöhen. Dazu gehörte zunächst die Integration einer automatischen Log-Erstellung in Form von Textdateien, um Fehlerquellen und Simulationsverläufe systematisch nachvollziehen zu können. Ergänzend wurde ein Soll-Ist-Vergleich implementiert, der die eingegebenen Produktionsaufträge den tatsächlich in der Simulation fertiggestellten Artikeln gegenüberstellt. Darüber hinaus wurde ein Excel-Export ergänzt, der alle nicht fertiggestellten Artikel mit zugehöriger Begründung ausweist. Zusätzlich wurde mit der Entwicklung einer Ausgabeliste begonnen, die die geplante Abarbeitungsreihenfolge der Aufträge visualisieren soll.

Im Zuge der Anwendung wurden zugleich technische und konzeptionelle Herausforderungen identifiziert. Auf technischer Ebene zeigte sich zunächst ein Stabilitätsproblem darin, dass Artikel trotz bereits überschrittener Deadline weiter produziert wurden, anstatt aus der Abarbeitung ausgeschlossen zu werden. Dies führte wiederholt zu fehlerhaften bzw. instabilen Simulationsläufen. Die zugrunde liegende Logik wurde daraufhin angepasst, sodass die Simulationen im weiteren Verlauf stabil und konsistent durchführbar waren. Auf konzeptioneller Ebene wurde deutlich, dass die ursprüngliche Nutzung des Tools zur detaillierten Reihenfolgeplanung in diesem Anwendungsumfeld nur eingeschränkt zweckmäßig war. Aufgrund der hohen Dynamik des betrachteten Produktionssystems und sehr kurzer Bearbeitungszeiten von etwa sieben Minuten je Produkt erwies sich eine minutengenaue Einplanung der Aufträge als wenig praktikabel.

Erzielte Ergebnisse

Der Use Case zeigte, dass das entwickelte Tool erfolgreich an reale Unternehmensdaten angebunden und für eine betriebsnahe simulationsgestützte Wochenplanung eingesetzt werden konnte. Insbesondere die automatisierte Übernahme realer ERP-Aufträge stellte einen wichtigen Schritt in Richtung praktischer Nutzbarkeit dar. Gleichzeitig wurden durch die ergänzten Logging- und Exportfunktionen die Transparenz, Nachvollziehbarkeit und Auswertbarkeit der Simulationsergebnisse deutlich verbessert.

Ein wesentliches Ergebnis bestand darin, dass sich der sinnvolle Anwendungsfokus des Tools in diesem Fall klar herausbildete. Während die detaillierte Sequenzplanung aufgrund der hohen Prozessdynamik nur begrenzt geeignet war, erwies sich die simulationsgestützte Vorausplanung des Materialbedarfs als praxisrelevanter und wirtschaftlich sinnvoller Einsatzbereich. Der Anwendungsfall bestätigte damit nicht nur die grundsätzliche Nutzbarkeit des entwickelten Tool-Kerns in einem realen industriellen Umfeld, sondern lieferte auch wichtige Erkenntnisse zur erforderlichen anwendungsspezifischen Fokussierung.

Darüber hinaus trug der Use Case wesentlich zur technischen Reifung des Tools bei. Die identifizierten Stabilitätsprobleme konnten behoben und zusätzliche Funktionen zur Fehleranalyse und Ergebnisaufbereitung integriert werden. Insgesamt stellt der Fall damit einen belastbaren Transfer- und Validierungsbeitrag dar, da das Tool nicht nur unter realen Bedingungen angewendet, sondern im Rahmen der Nutzung auch gezielt weiterentwickelt und an betriebliche Anforderungen angepasst wurde.

3.5 Arbeitspaket 5: Überführung der Methodik in ein IT-Werkzeug und vorhabenbegleitende Optimierung

Durchgeführte Arbeiten

Die Materialflusssimulation wurde als eine Python-basierte Standalone-Anwendung entwickelt, um eine kostengünstige und zugängliche Lösung für KMU zu schaffen. Kommerzielle Tools wie Plant Simulation oder AnyLogic kamen wegen ihrer hohen Lizenzkosten nicht in Frage. Kostenlose Anwendungen wie JaamSim konnten die spezifischen Anforderungen (Stabilität, Kompatibilität von Programmiersprachen) an die Integration von Reinforcement Learning nicht erfüllen. Zudem bietet Python eine hohe Flexibilität bei der Schnittstellenanbindung, eine breite Entwickler-Community sowie eine direkte Einbindung moderner Machine-Learning- und Simulationsbibliotheken, was die Weiterentwicklung der Lösung langfristig sichert. Die Entwicklung des Standalone-Tools erfolgte in der Programmiersprache Python mithilfe der GUI-Bibliothek PyQt5 und basierte auf den Anforderungen aus AP 1. Alle Bedienelemente und Interfaces können sowohl auf Deutsch als auch auf Englisch angezeigt werden.

Erzielte Ergebnisse

Das Software-Tool PrOPPlan ermöglicht die Verwaltung von Produktionsressourcen, Produktvorranggraphen und Auftragsdaten sowie die Konfiguration simulations- und steuerungsrelevanter Parameter. Außerdem bietet es dem Nutzer die Möglichkeit zur Durchführung von zufallsbasierten Simulationsläufen mit einer anschließenden statistischen Auswertung – dies ermöglicht eine Abschätzung von Optimierungspotenzialen im bestehenden Produktionssystem. Darüber hinaus bietet es Funktionen zur Durchführung von RL-Trainingsläufen und Experimenten sowie die Möglichkeit, gespeicherte Anwendungsszenarien im JSON-Format zu laden. Abbildung 7 demonstriert die Benutzeroberfläche von PrOPPlan nach dem Laden eines vorher gespeicherten Anwendungsszenarios:

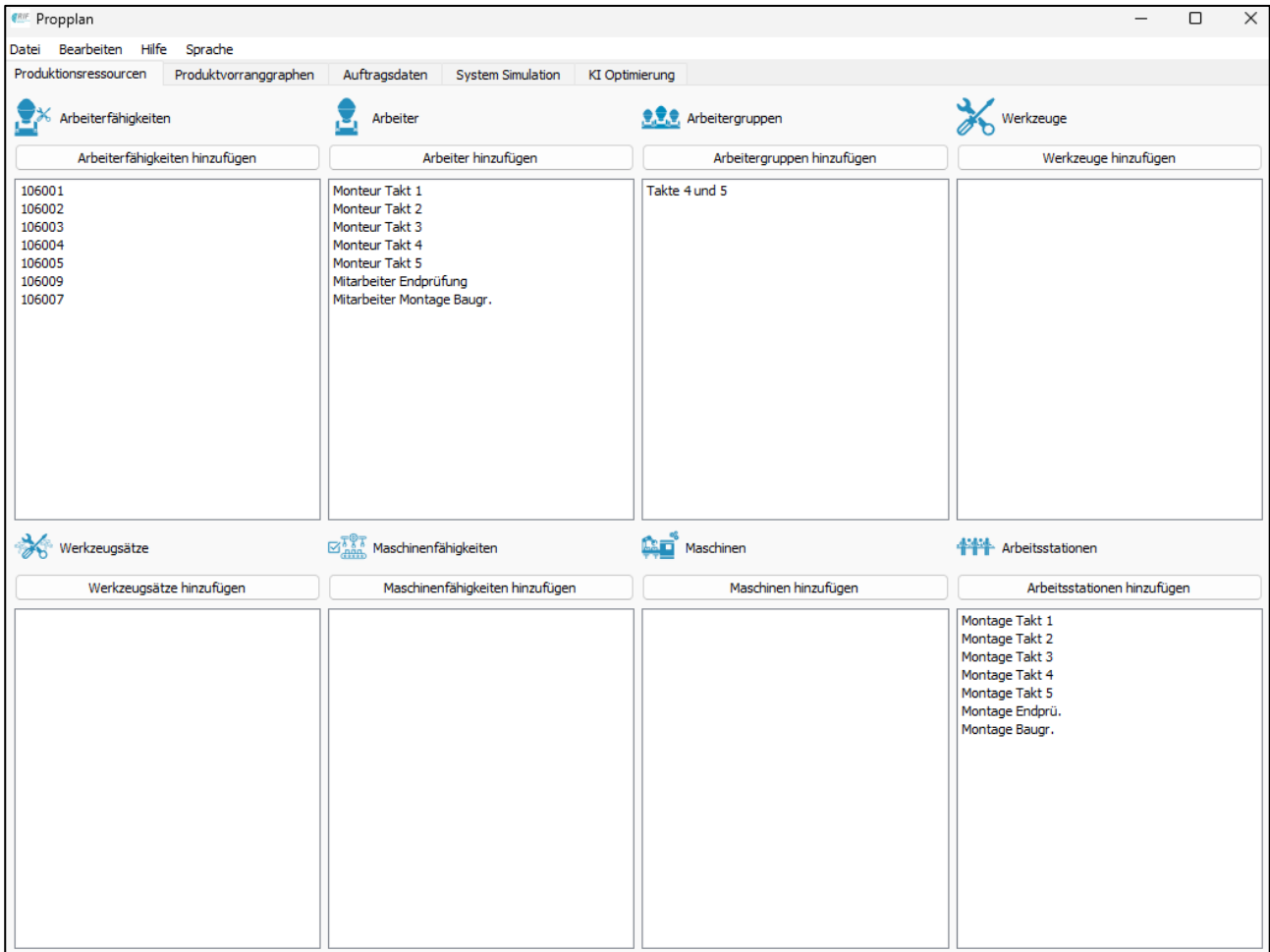


Abbildung 7: Tab 1 der Benutzeroberfläche von Propplan – Eingabe von Produktionsressourcen

Der erste Tab von ProOPPlan („Produktionsressourcen“) ermöglicht eine strukturierte Eingabe von acht der neun planungsrelevanten Entitäten gemäß dem in AP 2 erarbeiteten Datenmodell. Dazu gehören Mitarbeiterfähigkeiten, Mitarbeiter (mit Mehrfachfähigkeiten und Gruppenzugehörigkeiten), Werkzeuge und Maschinen, die jeweils mit spezifischen Eigenschaften wie Rüstzeiten und Leistungsparametern definiert werden können. Zusätzlich können Arbeitsstationen konfiguriert werden, indem Maschinen, Mitarbeiter, Werkzeuge und Puffer mit festen Zuordnungen oder flexiblen Auswahlmöglichkeiten hinterlegt werden. Eine Erweiterung basierend auf Unternehmensrückmeldungen erlaubt die Berücksichtigung dynamischer Zustände von Werkzeugen. Die Eingabemaske gewährleistet flexible Modellierung von Produktionssystemen. Wichtig hier zu erwähnen ist, dass die Eingabe vieler der hier aufgelisteten Parameter optional ist und sinngemäß Defaultwerte besitzt.

Produktvorranggraphen können im zweiten Tab der Anwendung eingegeben werden (Abbildung 8). Im Wesentlichen werden in dieser Eingabestufe Strukturstücklisten mit Arbeitsplänen der jeweiligen Produkte in anschaulicher Form vereinigt.

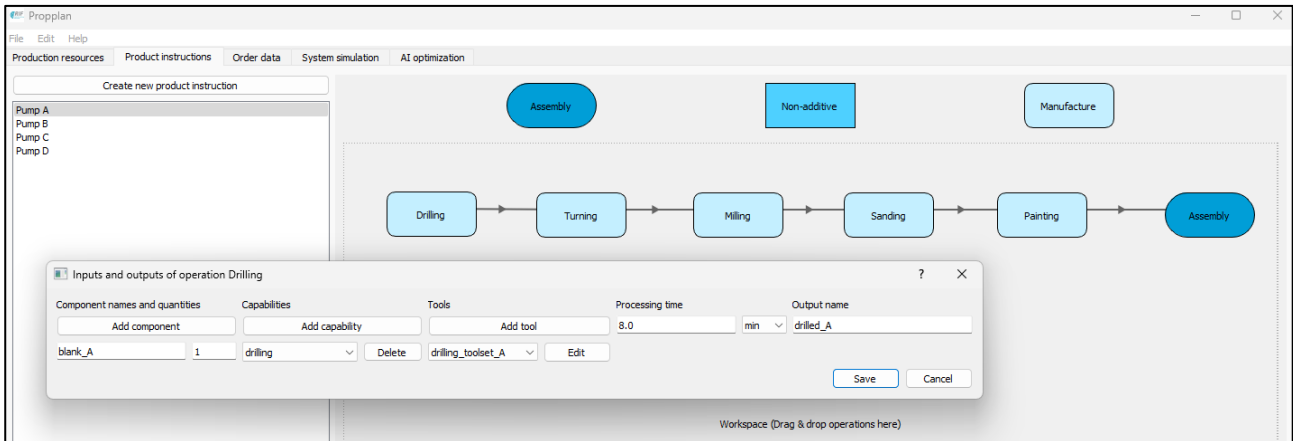


Abbildung 8: Eingabe von Produktvorranggraphen inklusive Operationsvoraussetzungen, Bearbeitungsdauer und Produkt

Für jedes Produkt müssen Stammdaten vorliegen, die den Ablauf und die Voraussetzungen der Operationen definieren. Benutzer können den Vorranggraphen per Drag-and-Drop erstellen und Operationen über Vorrangbeziehungen verknüpfen. Jede Operation verfügt über eine Eingabemaske zur Definition von Materialien, benötigten Fähigkeiten, Werkzeugen und Bearbeitungszeiten.

Zur besseren Visualisierung und Planung werden drei farb- und formkodierte Operationstypen genutzt, die unterschiedliche Parallelisierungs- und Chargenbildungsmöglichkeiten sowie materialflussbezogene Restriktionen kennzeichnen. Dies erleichtert die ergonomische Darstellung und hebt den wertschöpfenden Beitrag der jeweiligen Operation hervor.

Dem formalisierten Modell von Produktionssystemen weiter folgend, befinden sich im dritten Tab die Auftragsdaten. Jeder Auftrag besitzt eine Produktliste mit bestellten Mengen, den frühesten Bearbeitungszeitpunkt (Freigabe) und den spätesten Fertigstellungspunkt (Liefertermin). Der Freigabezeitpunkt kann dabei entweder aus einer vorgelagerten Grobplanung stammen oder zu Simulationszwecken als Zeitpunkt der Kundenbestellung interpretiert werden.

Im vierten Tab kann eine beliebige Anzahl von Experimenten simuliert werden, die auf zufallsbasierten Entscheidungen zur Operationsauswahl (Sequencing) und zur Operationszuteilung (Routing) beruhen (vgl. Abbildung 9). Die daraus resultierenden prozessbezogenen Ergebnisse werden hinsichtlich verschiedener Kenngrößen analysiert, wie beispielsweise der durchschnittlichen Mitarbeiterauslastung oder der Lieferzeit. Die Ergebnisse werden entsprechend aufbereitet und zur weiteren Auswertung bereitgestellt. Der Nutzer kann anschließend festlegen, nach welchen Kennzahlen die Ergebnisse sortiert werden sollen, und erhält darauf basierend eine Visualisierung des Arbeitsplans, der Ressourcenauslastungen oder der Pufferverläufe in Abhängigkeit von der Zeit. Zusätzlich werden stochastische Kennzahlen ausgewertet und deren Minimal-, Maximal- und Durchschnittswerte angezeigt. Darüber hinaus können in diesem Fenster auch zuvor mittels Reinforcement Learning erzeugte Produktionspläne geladen und visualisiert werden.

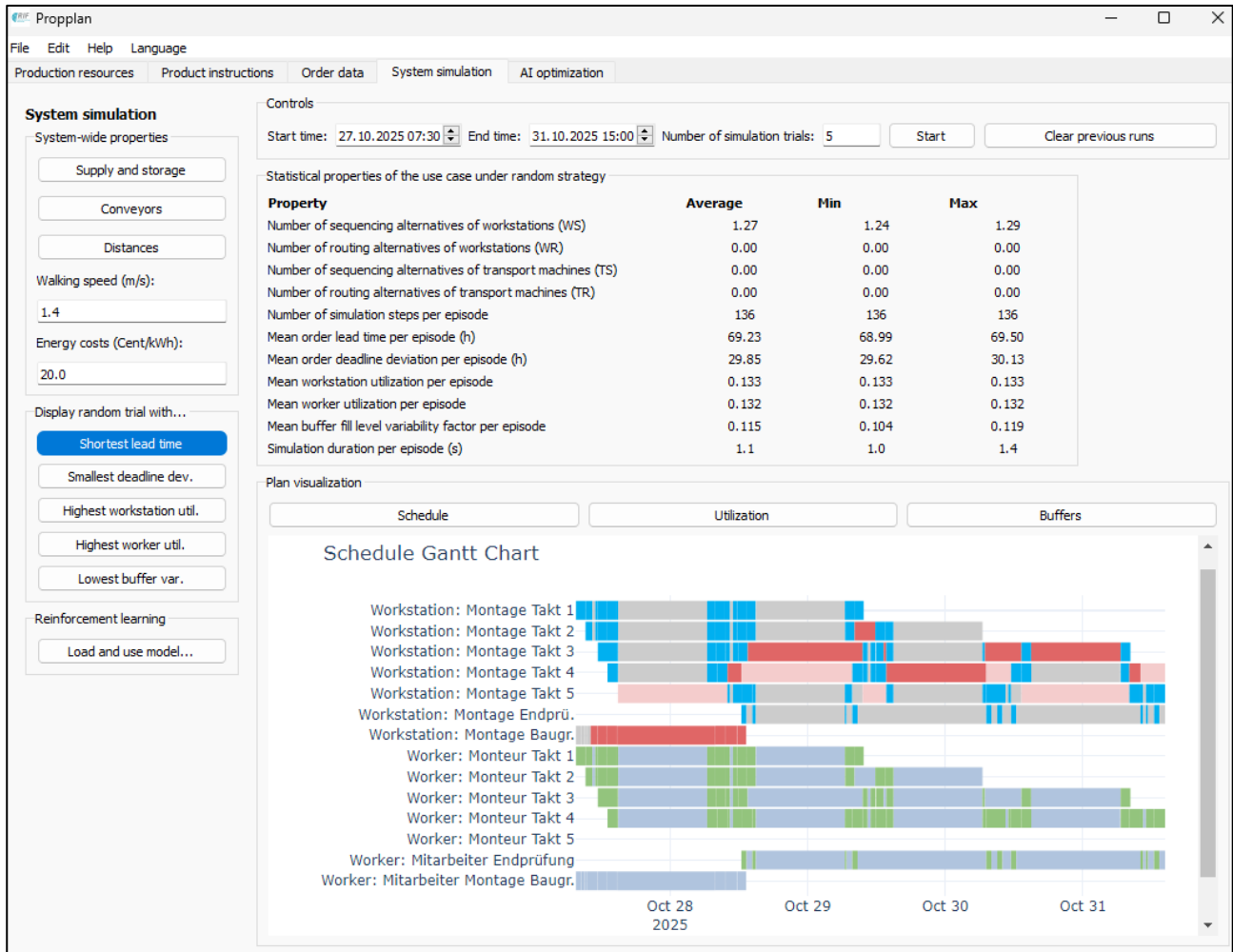


Abbildung 9: Simulationsdarstellung eines realen Produktionssystems mit dem Arbeitsplan, welcher die kürzeste Durchlaufzeit hat

Unter Verwendung der nun vollständig erstellten Simulation können Nutzer im fünften Tab eigene Optimierungsexperimente konfigurieren und ausführen. Dabei besteht die Möglichkeit, die einzelnen Versuchsparameter gezielt zu variieren, um für den jeweiligen Anwendungsfall verbesserte Planungsergebnisse zu erzielen. Gleichzeitig stehen vordefinierte Parameterwerte zur Verfügung, mit denen ein Reinforcement-Learning-basierter Optimierungsprozess unmittelbar gestartet werden kann, sodass keine vertieften Vorkenntnisse im Bereich Reinforcement Learning erforderlich sind. Abbildungen 10-13 zeigen exemplarische Konfigurationsdialoge, deren Ausfüllen den Start eines Simulations- oder Trainingslaufs mit dem virtuellen Produktionssystem ermöglicht. Nach Abschluss eines Laufs können erzeugte Produktionspläne, zeitliche Verläufe von Ressourcenauslastungen und Pufferständen sowie relevante Kennzahlen interaktiv analysiert werden. Zusätzlich ist im fünften Tab eine tabellarische Übersicht der Experimentergebnisse verfügbar.

Abbildung 10: Konfiguration von RL-MuZero

Raw state variables		
Data	Number of entries	Observable
Workers: location (workstation)	176	<input checked="" type="checkbox"/>
Workers: location (transport)	66	<input type="checkbox"/>
Workers: destination (workstation)	176	<input type="checkbox"/>
Workers: destination (transport)	66	<input type="checkbox"/>
Workers: status	44	<input checked="" type="checkbox"/>
Tools: location	384	<input type="checkbox"/>
Tools: status	72	<input type="checkbox"/>
Workstations: status	176	<input checked="" type="checkbox"/>
Input buffers: fill level	16	<input type="checkbox"/>

Aggregated state variables		
Data	Number of entries	Observable
Buffers: average fill level	32	<input checked="" type="checkbox"/>
Buffers: fill level variability	32	<input type="checkbox"/>
Workstations: productive time ratio	16	<input type="checkbox"/>
Workstations: setup time ratio	16	<input checked="" type="checkbox"/>
Workers: productive time ratio	11	<input type="checkbox"/>
Workers: setup time ratio	11	<input checked="" type="checkbox"/>
Workers: walking time ratio	11	<input type="checkbox"/>

Observation space size
Total number of vector entries: **463**

Abbildung 11: Konfiguration des Beobachtungsraums

Decision type	Direct	Indirect
Workstation routing	<input checked="" type="checkbox"/>	
Workstation sequencing	<input checked="" type="checkbox"/>	
Transport routing	<input type="checkbox"/>	Least queued transport orders (LQTO)
Transport sequencing	<input type="checkbox"/>	FIFO

Abbildung 12: Konfiguration des Aktionsraums

Goal	KPI	Scale to 1 point	Unit
Punish	Mean order lead time	1	h
Punish	Mean absolute order deadline deviation	1	h
Ignore	Mean productive time ratio of workstations	e.g., 1.0	
Reward	Mean productive time ratio of workers	0.05	
Ignore	Mean buffer fill variability factor	e.g., 1.0	

Abbildung 13: Konfiguration der Belohnungsfunktion

Neben der umfassenden Programmierung der Simulationslogik in Python wurden Anregungen von industriellen Partnern berücksichtigt und Funktionen mit Eingabefeldern nach ihrem Wunsch umgesetzt. Dazu zählen unter anderem Eingabemöglichkeiten für Anwesenheitszeiten der Mitarbeiter (Schmitz Medical GmbH), Detaillierung von Werkzeugen und Hilfsmitteln (Polymerge und QTS) und einblendbare Erklärttexte/Tooltips (Jung).

Wie im Förderantrag vorgesehen, wurde im Rahmen von AP 5 ein IT-Werkzeug zur nutzerfreundlichen Verwendung der selbstoptimierenden MFS entwickelt und der breiten Öffentlichkeit zur Verfügung gestellt (<https://github.com/institute-of-production-systems/reinforcement-learning-in-production-planning>).

Zentrale praxisrelevante Funktionen der PROPlan-Anwendung lassen sich wie folgt zusammenfassen:

1. Eine flexible Modellierung von Produktionssystemen unter Berücksichtigung zahlreicher praxisnaher Aspekte und betrieblicher Logiken, die ein fundiertes Verständnis von Planungsproblemen unterstützt.

2. Die Möglichkeit zur KI-gestützten, regelbasierten sowie manuellen Planung innerhalb eines einheitlichen Workflows, wodurch sowohl eine effiziente Einarbeitung als auch eine systematische Vergleichbarkeit unterschiedlicher Planungsansätze gewährleistet wird.
3. Die explizite Berücksichtigung der in der industriellen Praxis häufig eingeschränkten Beobachtbarkeit von Produktionssystemen, sodass der Planungsalgorithmus auch unter realistischen Informationsbedingungen anwendbar bleibt.
4. Die Anpassbarkeit des Optimierungsziels durch eine konfigurierbare Belohnungsfunktion, die eine zielgerichtete Ausrichtung der Optimierung erlaubt.

Die vorgestellte Benutzeroberfläche ermöglicht es insbesondere Anwendern in kleinen und mittleren Unternehmen, Reinforcement Learning auf ihre jeweiligen Produktionssysteme anzuwenden und diese gemäß individuell definierter Zielsetzungen zu optimieren. Die Nutzung vordefinierter Parameterwerte sowie kontextsensitiver Warnhinweise bei unvollständigen Eingaben erleichtert den Einstieg und reduziert den Konfigurationsaufwand. Darüber hinaus bietet der vierte Tab die Möglichkeit, das modellierte Produktionssystem im Hinblick auf seine grundsätzliche Eignung für eine Optimierung mittels Reinforcement Learning zu evaluieren. Im letzten Tab wird der Nutzer schließlich in die Lage versetzt, Trainingsparameter gezielt anzupassen und damit Einfluss auf die resultierenden Kennzahlen zu nehmen.

3.6 Arbeitspaket 6: Projektmanagement, Öffentlichkeitsarbeit, Dokumentation und Ergebnistransfer

Durchgeführte Arbeiten

Die Dokumentation und Dissemination der Ergebnisse erfolgten parallel zur Bearbeitung der AP 1 bis 5. Zusätzlich zur Koordination der Arbeitstreffen und Untersuchungen im PA wurde ein Ergebnistransfer in Industrie und Forschung angestrebt. Dies beinhaltete die diskriminierungsfreie Bereitstellung der (Teil-)Ergebnisse in einer breiten Fachöffentlichkeit. Angestrebte Veröffentlichungsformen waren neben dem Abschlussbericht, weitere Maßnahmen im Sinne einer Öffentlichkeitsarbeit, wie bspw. praxisorientierte Workshops beim Netzwerk Industrie Ruhr-Ost (NIRO e.V.), die Ergebnisveröffentlichung in (inter-)nationalen wissenschaftlichen Tagungsbänden und im Internet, sowie die Vorstellung der Inhalte auf Fachtagungen und Konferenzen.

Erzielte Ergebnisse

Es wurden im Berichtszeitraum neben zahlreichen bilateralen Arbeitstreffen, vier Sitzungen des projektbegleitenden Ausschusses (sowohl digital als auch in Präsenz) und ein anwenderorientierter Workshop in den Räumlichkeiten von NIRO e.V. durchgeführt. Der Zwischenstand der Tool-Entwicklung wurde stets auf GitHub veröffentlicht. Darüber hinaus wurden die Forschungs- und Entwicklungsergebnisse am 25.09.2025 in einem Vortrag auf der ASIM-Fachtagung *Simulation in Produktion und Logistik 2025* in Dresden vorgestellt. Bibliographische Angaben zur Veröffentlichung von Polikarpov et al. (2025) sind dem Literaturverzeichnis zu entnehmen.

4 Verwendung der Zuwendung

Die beantragbaren Zuwendungen werden in Zuwendungen für wissenschaftlich-technisches Personal, Geräte und Leistungen Dritter untergliedert. Im Rahmen des Forschungsprojektes PrOPPlan beschränkten sich die Zuwendungen ausschließlich auf wissenschaftlich-technisches Personal. Auf die Beschaffung von Geräten sowie auf Leistungen Dritter wurde gemäß Forschungsantrag verzichtet.

Wissenschaftlich-technisches Personal (Einzelansatz A.1 des Finanzierungsplans):

Die Erarbeitung der Projektergebnisse erfolgte durch Personal mit abgeschlossener wissenschaftlicher Ausbildung (HPA-A) sowie durch wissenschaftliche Hilfskräfte (HPA-F). Entsprechend dem im Arbeitsplan definierten Aufgabenspektrum wurde das Forschungsvorhaben – unter Berücksichtigung der bewilligten kostenneutralen Laufzeitverlängerung vom 30.06.2025 bis zum 31.12.2025 – in einem Gesamtzeitraum von 30 Monaten bearbeitet. Das Projekt wurde gemäß der Arbeitsplanung im Projektantrag am 01.07.2023 aufgenommen.

Die im Antrag dargelegte methodische Vorgehensweise erwies sich im Rahmen der Durchführung als zielführend, um eine systematische Erreichung der Forschungsziele zu gewährleisten. Der für die Bearbeitung erforderliche personelle Aufwand der Forschungsstelle RIF Institut für Forschung und Transfer e.V. ist in Tabelle 1 detailliert aufgeführt und beläuft sich auf insgesamt 29 Personenmonate (PM).

Tabelle 1: Eingesetzte personelle Ressourcen im Berichtszeitraum

Personelle Ressourcen	Eingesetzte personelle Ressourcen [PM] im Berichtszeitraum		
	2023	2024	2025
HPA-A, FE 1 (PM)	0,00	7,25	21,75
HPA-F, FE 1 (PM)	0,00	7,00	4,75
Gesamt/Berichtszeitraum (HPA-A)	29,00		
Gesamt/Berichtszeitraum (HPA-F)	11,75		

Die fachliche Bearbeitung des Forschungsvorhabens durch das RIF Institut für Forschung und Transfer e.V. erforderte über die gesamte Projektlaufzeit den Einsatz von wissenschaftlichem Personal mit hochspezialisierten Fachkenntnissen in der Produktionstechnik, der Datenanalyse sowie der Softwareentwicklung. Zur wissenschaftlichen Umsetzung der im Arbeitsprogramm definierten Arbeitspakete wurden zwei wissenschaftliche Mitarbeitende (HPA-A) mit einem Beschäftigungsumfang von insgesamt 29 Personenmonaten eingesetzt, die durch wissenschaftliche Hilfskräfte (HPA-F) unterstützt wurden.

Ein wissenschaftlicher Mitarbeiter mit produktions- und datenanalytischem Forschungshintergrund führte das Vorhaben über die gesamte Laufzeit federführend durch. Dabei floss insbesondere Expertenwissen aus dem Bereich der operativen Produktionsplanung und -steuerung in die wissenschaftliche Bearbeitung der beantragten Arbeitspakete ein. Neben der inhaltlichen Ausarbeitung oblag dieser Stelle das übergeordnete Projektmanagement sowie die Koordination und Organisation der regelmäßigen Arbeitstreffen des

projektbegleitenden Ausschusses (PA). Ein zweiter wissenschaftlicher Mitarbeiter mit spezifischen Programmierkenntnissen forcierte insbesondere in der zweiten Projekthälfte die informationstechnische Umsetzung des IT-Werkzeugs. Hierbei wurde Reinforcement Learning als Disziplin des Maschinellen Lernens implementiert, um die im Projektantrag beschriebenen Iterationsstufen des Werkzeugs technologisch zu realisieren.

Durch einen kontinuierlichen und engen fachlichen Austausch zwischen der federführenden Projektleitung und der IT-Entwicklung wurde sichergestellt, dass die Konzepte der Produktionssteuerung und die Algorithmen des Maschinellen Lernens passgenau in die Softwareentwicklung einfließen. Das wissenschaftliche Personal wurde bei Routinearbeiten und wiederholenden Tätigkeiten durch Hilfskräfte unterstützt, welche insbesondere Aufgaben in der Literaturrecherche, der Datenpflege sowie der Vor- und Nachbereitung der Analyseergebnisse nach Weisung übernahmen. Die im Rahmen des Projekts bewilligten finanziellen Mittel wurden unter strikter Einhaltung der Grundsätze der Wirtschaftlichkeit und Sparsamkeit sowie im Einklang mit dem genehmigten Gesamtfinanzierungsplan zweckentsprechend zur Erreichung der Projektziele eingesetzt.

Geräte (Einzelansatz B des Finanzierungsplans)

Im Zuge der Erarbeitung der Projektergebnisse konnte auf die Anschaffung von Gerätschaften verzichtet werden.

Leistungen Dritter (Einzelansatz C des Finanzierungsplans)

Im Zuge der Erarbeitung der Projektergebnisse konnte auf Leistungen Dritter verzichtet werden. Unentgeltlich erbrachte Leistungen im Sinne von vAWs sind dem zahlenmäßigen Nachweis zu entnehmen.

5 Notwendigkeit und Angemessenheit der geleisteten Arbeit

Die Realisierung der angestrebten Forschungsziele im Projekt PrOPPlan basierte maßgeblich auf einer engen Verzahnung von wissenschaftlicher Erarbeitung und industrieller Validierung. Die inhaltliche Umsetzung erfolgte entsprechend den Vorgaben des bewilligten Antrags durch qualifiziertes wissenschaftliches Personal unter Einbindung studentischer Hilfskräfte. Ein wesentliches Element zur Sicherstellung der Notwendigkeit und Angemessenheit der methodischen Qualitätssicherung bildete der diskursive Austausch im Rahmen des projektbegleitenden Ausschusses (PA), wodurch branchenspezifische Anforderungen und praktisches Expertenwissen unmittelbar in die Gestaltung der operativen Produktionsprogrammplanung einfließen konnten. Die strategische Zusammensetzung des PA aus industriellen Anwenderunternehmen, Systementwicklern und Simulationsberatern bildete dabei das Fundament für die technische Belastbarkeit und die universelle Einsetzbarkeit der entwickelten Konzepte, insbesondere im Kontext des Reinforcement Learning.

Trotz punktueller zeitlicher Abweichungen vom ursprünglichen Meilensteinplan konnte die hohe Qualität der Projektergebnisse über die gesamte Laufzeit gewahrt werden. Diese Modifikationen resultierten einerseits aus personellen Vakanz in der Startphase und andererseits aus fachspezifischen Impulsen der PA-Mitglieder. Um den Projektfortschritt angemessen zu fördern, wurde die softwaretechnische Umsetzung (AP 5) bereits mehrere Monate früher als geplant initiiert, da eine funktionsfähige IT-Basis für die erfolgreiche Bearbeitung der Arbeitspakete 3 und 4 unabdingbar war. In diesem Zusammenhang erwies es sich innerhalb der Simulationsmethodik als wissenschaftlich geboten, die Modellbildung dynamisch an den Kenntnisstand der späteren Anwender anzupassen, anstatt einem statischen Schema zu folgen. Der daraus resultierende inhaltliche Mehraufwand war notwendig, um eine gesteigerte Robustheit und Praxisnähe des Systems zu

kompensieren. Zur finalen Absicherung dieser erweiterten Entwicklungsumfänge wurde eine kostenneutrale Laufzeitverlängerung bis zum 31.12.2025 in Anspruch genommen.

Die Angemessenheit der Entwicklungen wurde zudem durch umfangreiche vorhabenbezogene Aufwendungen der beteiligten Wirtschaftspartner validiert. Neben den vier zentralen Sitzungen des projektbegleitenden Ausschusses (PA1 bis PA4) wurden zahlreiche bilaterale Termine und Vor-Ort-Besuche bei Partnerunternehmen durchgeführt, um Datenbedarfe und Systemanforderungen für die selbstoptimierende Materialflusssimulation zu verifizieren. Besondere Schwerpunkte bildeten hierbei fachspezifische Workshops zur Produktionsprogrammplanung, Termine bei Netzwerkpartnern sowie dedizierte Experteninterviews und Rücksprachen zur Modellentwicklung und bestehenden -anforderungen. Diese aktive Partizipation der Industriepartner gewährleistete eine kontinuierliche Spiegelung der Forschungsergebnisse an der betrieblichen Praxis.

Mit dem Ende des Bewilligungszeitraums sind die angestrebten Forschungsziele vollständig realisiert worden. Durch die tiefgreifende Partizipation der industriellen Anwender wurden zudem strategische Ansatzpunkte für künftige Erweiterungen des Verfahrens definiert, welche die Automatisierungspotenziale in der Produktionsprogrammplanung weiter ausschöpfen können. Neben dem unmittelbaren Nutzen für die Praxis wurde durch die Integration von Forschungsfragen in studentische Abschlussarbeiten ein nachhaltiger Transfer der Projektergebnisse in die universitäre Lehre sichergestellt.

6 Darstellung des wissenschaftlich-technischen und wirtschaftlichen Nutzens der erzielten Ergebnisse insbesondere für KMU sowie ihres innovativen Beitrags und ihrer industriellen Anwendungsmöglichkeiten

Wissenschaftlich-technischer Nutzen

Das Forschungsvorhaben PrOPPlan leistet einen wesentlichen wissenschaftlich-technischen Beitrag zur operativen Produktionsprogrammplanung, indem eine methodische und softwaretechnische Verbindung zwischen ereignisdiskreter Materialflusssimulation und Reinforcement Learning geschaffen wird. Zentrales Ergebnis ist die Entwicklung eines formal konsistenten Simulations- und Entscheidungsrahmens, der es erlaubt, komplexe Produktionssysteme mit realitätsnahen Restriktionen (z. B. Vorranggraphen, begrenzte Ressourcen, Mitarbeiterqualifikationen, Rüstvorgänge und partielle Beobachtbarkeit) für selbstlernende Optimierungsverfahren zugänglich zu machen.

Besonders hervorzuheben ist die explizite Berücksichtigung unvollständiger Systeminformationen im Sinne partieller Beobachtbarkeit. Während viele bestehende Forschungsarbeiten implizit von vollständiger Zustandskenntnis ausgehen, adressiert PrOPPlan gezielt praxisnahe Informationsdefizite, wie sie insbesondere in KMU vorliegen. Die entwickelte Zustandsraumformalisierung erlaubt es, unterschiedliche Grade der Beobachtbarkeit systematisch zu untersuchen und deren Einfluss auf die Leistungsfähigkeit lernbasierter Planungsansätze zu analysieren.

Darüber hinaus stellt das Projekt einen methodischen Beitrag zur integrierten Betrachtung mehrerer operativer Planungsaufgaben dar. Anstelle isolierter Optimierung einzelner Teilprobleme (z. B. reine Reihenfolgeplanung) ermöglicht PrOPPlan die simultane Berücksichtigung von Losgrößenbildung, Feinterminierung, Reihenfolgeplanung und Auftragsfreigabe innerhalb eines einheitlichen Entscheidungsmodells.

Wirtschaftlicher Nutzen, insbesondere für KMU

Der wirtschaftliche Nutzen von PrOPPlan liegt insbesondere in der deutlichen Absenkung von Eintrittsbarrieren für den Einsatz simulations- und KI-gestützter Produktionsplanung in kleinen und mittleren Unternehmen. Kommerzielle Simulations- und Optimierungswerkzeuge sind für KMU häufig aufgrund hoher Lizenzkosten, eingeschränkter Anpassbarkeit und des erforderlichen Spezialwissens nicht praktikabel. PrOPPlan stellt demgegenüber ein frei verfügbares, quelloffenes IT-Werkzeug bereit, das ohne zusätzliche Lizenzkosten eingesetzt und schrittweise an die jeweiligen betrieblichen Gegebenheiten angepasst werden kann.

Durch die Möglichkeit, operative Produktionsentscheidungen datenbasiert und simulationsgestützt zu analysieren, können KMU Planungsentscheidungen transparenter gestalten und manuelle Dispositionsaufwände reduzieren. Insbesondere in Produktionsumgebungen mit hoher Variantenvielfalt, schwankender Auftragslage oder begrenzten personellen Ressourcen unterstützt PrOPPlan eine robustere und nachvollziehbare Entscheidungsfindung.

Ein weiterer wirtschaftlicher Mehrwert besteht darin, dass das Tool sowohl für einfache simulationsbasierte Analysen als auch für weiterführende KI-gestützte Optimierung genutzt werden kann. Unternehmen können somit ohne unmittelbaren Einsatz von Reinforcement Learning zunächst ein besseres Systemverständnis aufbauen und Optimierungspotenziale identifizieren, bevor sie schrittweise lernbasierte Verfahren einsetzen. Dies reduziert Investitionsrisiken und erleichtert die praktische Einführung neuer Planungsmethoden.

Innovativer Beitrag

Der innovative Kern von PrOPPlan liegt in der konsequenten Verbindung von selbstoptimierender Materialflusssimulation mit einem hochgradig konfigurierbaren Reinforcement-Learning-Rahmen für die operative Produktionsprogrammplanung. Im Gegensatz zu vielen bestehenden Ansätzen, die auf feste Heuristiken oder stark problem-spezifische RL-Implementierungen beschränkt sind, erlaubt PrOPPlan eine flexible Definition von Beobachtungsraum, Aktionsraum und Belohnungsfunktion.

Diese Konfigurierbarkeit ermöglicht es, unterschiedliche Planungsstrategien, Informationsannahmen und Zielsysteme systematisch zu vergleichen. Dadurch wird Reinforcement Learning als nachvollziehbares Entscheidungsinstrument nutzbar gemacht. Besonders innovativ ist dabei die Möglichkeit, auch indirekte Steuerungsansätze (z. B. Auswahl von Prioritätsregeln) innerhalb desselben Frameworks zu untersuchen und mit direkt steuernden RL-Ansätzen zu vergleichen.

Zudem stellt die entwickelte Methodik eine Abkehr von streng sequenziellen Planungslogiken dar. Durch die integrierte Betrachtung mehrerer operativer Entscheidungen in einem kohärenten Modell können Zielkonflikte explizit berücksichtigt und globale Planungslösungen angestrebt werden, anstatt lokal optimale Einzelösungen zu erzeugen.

Industrielle Anwendungsmöglichkeiten

Die im Projekt entwickelten Ergebnisse sind in einer Vielzahl industrieller Anwendungsszenarien einsetzbar, insbesondere in Produktionssystemen mit komplexen Arbeitsplänen, variabler Ressourcenausstattung und hoher Änderungsdynamik. Typische Anwendungsfelder sind unter anderem der Maschinen- und Anlagenbau, die Variantenfertigung sowie Montage- und Werkstattfertigungen.

Das IT-Werkzeug PrOPPlan kann sowohl zur operativen Feinplanung als auch zur simulationsgestützten Bewertung alternativer Produktionsprogramme und Steuerungsstrategien eingesetzt werden. Durch die

Nutzung realitätsnaher Eingangsdaten und die Möglichkeit zur Abbildung betrieblicher Restriktionen eignet sich das Tool sowohl für explorative Analysen als auch für die Vorbereitung konkreter Planungsentscheidungen.

Die im Projekt durchgeführte Anwendung und Abstimmung mit industriellen Partnern, unter anderem anhand realer Produktionsdaten, unterstreicht die grundsätzliche Praxistauglichkeit des Ansatzes. Gleichzeitig bietet die offene Architektur des Werkzeugs eine geeignete Basis für zukünftige Erweiterungen, etwa zur Kopplung mit bestehenden IT-Systemen oder zur weiteren Automatisierung operativer Entscheidungsprozesse.

7 Wissenstransfer in die Wirtschaft

Der Wissens- und Ergebnistransfer in die Wirtschaft erfolgte im Projektverlauf auf mehreren Ebenen. Zentrale Plattform hierfür waren die regelmäßigen Sitzungen des Projektbegleitenden Ausschusses (PA), in denen Zwischenergebnisse vorgestellt, methodische Ansätze diskutiert und praxisrelevante Anforderungen validiert wurden. Die kontinuierliche Einbindung der Unternehmensvertreter stellte sicher, dass die entwickelten Methoden und Werkzeuge frühzeitig an reale industrielle Rahmenbedingungen angepasst wurden.

Über die formalen PA-Sitzungen hinaus ergaben sich weiterführende Transferimpulse durch einen Workshop im Unternehmensnetzwerk NIRO zum Thema „KI in der Produktionsplanung und -steuerung“. In diesem Rahmen wurden neben anderen Ansätzen zur Nutzung von KI in der PPS zentrale Projektergebnisse sowie das zugrunde liegende Konzept der selbstoptimierenden MFS vorgestellt und mit Unternehmen außerhalb des ursprünglichen PA-Kreises diskutiert. Das Interesse an den vorgestellten Ansätzen führte dazu, dass weitere Unternehmen konkrete Anwendungsfälle identifizierten und sich aktiv in die Weiterentwicklung einbrachten.

In der Folge konnten zwei zentrale Use Cases initiiert und begleitet werden. Dabei handelte es sich um ein kleineres mittelständisches Unternehmen sowie ein größeres Industrieunternehmen mit deutlich unterschiedlichen Produktionsstrukturen und organisatorischen Voraussetzungen. Die methodische Übertragbarkeit auf diese unterschiedlichen Kontexte unterstreicht die Generalisierbarkeit des im Projekt entwickelten Ansatzes.

Der Transfer erfolgte begleitend, d. h. das entwickelte Tool wurde durch Mitarbeiter der Unternehmen sowie die Projektmitarbeiter des RIF angewandt und weiterentwickelt. So konnte eine direkte Verbindung in das ERP-System eines Anwenders realisiert werden. Dadurch konnte sowohl die Anwendbarkeit unter realen Bedingungen validiert als auch wertvolles Feedback für die iterative Verbesserung des Werkzeugs gewonnen werden.

Ein weiterer wesentlicher Transfermechanismus besteht in der freien Verfügbarkeit des entwickelten Tools für Forschungs- und Demonstrationszwecke. Die Ergebnisse sind so aufbereitet, dass sie von Unternehmen eigenständig genutzt und in bestehende Planungsprozesse integriert werden können. Das Werkzeug bildet die Grundlage für weitere Forschungs- und Transferprojekte im Rahmen der reaktiven, kurzfristigen Planung.

Insgesamt zeigt sich, dass der Wissenstransfer nicht ausschließlich über formale Berichts- und Ausschussstrukturen erfolgte, sondern maßgeblich durch aktive Netzwerkveranstaltungen, praxisnahe Demonstrationen und konkrete Anwendungsfälle in Unternehmen getragen wurde. Die Kombination aus strukturiertem PA-Dialog, öffentlicher Netzwerkveranstaltung, dem Leitfaden und realen Use Cases hat maßgeblich zur nachhaltigen Verankerung der Projektergebnisse in der Wirtschaft beigetragen.

8 Durchgeführte Transfermaßnahmen

Die konkreten Maßnahmen zur Überführung der Projektergebnisse in die Wirtschaft sind in Tabelle 2 zusammenfassend dargestellt.

Tabelle 2: Durchgeführte Transfermaßnahmen während der Projektlaufzeit

Transfermaßnahme	Übergeordnetes Ziel	Organis. Rahmen	Zeitpunkt
Projektinformation im Internet über den aktuellen Stand des Vorhabens	<ul style="list-style-type: none"> • Informationen zu den Projektergebnissen • Verbreitung der Projektidee an potentielle Anwender 	www.ips.do , www.bvl.de	Fortlaufend ab Projektbeginn (AP7)
Sitzungen des projektbegleitenden Ausschusses zum Meilenabschluss	<ul style="list-style-type: none"> • Diskussion und Abstimmung des Projektvorgehens • Sicherstellung von Akzeptanz und Nutzen für den PA 	Vier Meilensteintreffen des PA im Projekt	Mitte 2024, 16.05.2025, 19.11.2025
Transfer in die Industrie durch Netzwerke und Verbände	<ul style="list-style-type: none"> • Ansprache von und werbewirksame Verbreitung an potentiell interessierte Unternehmen und Anwender 	Mithilfe der Netzwerkpartner im PA	Fortlaufende Kommunikation s.o.
Präsentation auf Wissensplattformen oder in Arbeitskreisen	Austausch und Ergebnistransfer in die (regionale) Wirtschaft durch Plattformen wie bspw. NIRO-Wissen	Fortlaufend	Fortlaufend in verschiedenen Terminen (NIRO etc.)
Studentische Abschlussarbeiten, wie Master- / Bachelor- oder Projektarbeiten im Vorhaben	Ausschreibung wissenschaftlicher Arbeiten zur Vermittlung wissenschaftl. Erkenntnisse im Themenfeld und zur Heranführung Studierender an selbständiges Arbeiten	Moodle-Lernplattform der TU Dortmund	Kooperation mit zwei Unternehmen und Abschlussarbeitenden
Transfer in die Wissenschaft durch Ergebnisvorstellung auf Tagungen, Konferenzen oder Kongressen	Schriftlicher Bericht und Vortrag der Projektergebnisse: ASIM 2024, 2025	Tages- oder Mehrtagesveranstaltungen, jeweils jährlich	ab Q4/2023
Exemplarische Einführung des entwickelten IT-Werkzeugs	Zeitgleiche Verifizierung und Validierung der Forschungsergebnisse in den Anwenderbetrieben	In Workshops	Ab 2025

Veröffentlichungen in nationalen und internationalen Fachzeitschriften oder Journalen	Wissens. Veröffentlichungen in Fachzeitschriften: Journalbeitrag: Simulation Notes Europe 2026 (in Begutachtung)	Fortlaufend, jedoch überwiegend zum Ende des Vorhabens	Q2/2026
Abschlussbericht	Verbreitung der Projektergebnisse	BVL, Eldorado	ab Q2/2026

9 Geplante spezifische Transfermaßnahmen nach der Projektlaufzeit

Tabelle 3 veranschaulicht die geplanten spezifischen Transfermaßnahmen nach Projektende inklusive des angestrebten Zeithorizontes.

Tabelle 3: Geplante Transfermaßnahmen nach Projektende

Maßnahme	Ziel	Ort/Rahmen	Termin
Akquise potenziell interessierter Unternehmen	Unmittelbarer Ergebnistransfer in die Industrie	Vorstellungen der Projektergebnisse bei Unternehmen und Verbänden	Fortlaufend (und auf Anfrage)
Schulungsunterlagen	Praxisnahe Aufbereitung der Ergebnisse zur Durchführung von Workshops	Angebot von Workshops zu „KI in der Produktionsprogrammplanung“ und Überführung in den NIRO KI Navigator	ab Q3/2025
Transfer in die Industrie durch Netzwerke und Verbände	Wissensch. didaktische Aufbereitung der Ergebnisse zur Einbindung in die universitäre Lehre	Ergebnistransfer in die TU-Veranstaltung „Industrial Data Science 1 & 2“	Wintersemester 26/27
Präsentation auf Wissensplattformen oder in Arbeitskreisen	Wissensch. Qualifikation des Bearbeiters des Forschungsvorhabens	Dissertation	In Arbeit
Dissertation und wissenschaftl. Publikationen			
Akquisition möglicher Folgeprojekte	Weiterentwicklung	IGF, ZIM, KMU Innovativ	ab Q3/2026
Publikation der Projektergebnisse im Internet	Informationstransfer und Austausch mit Interessenten aus Industrie und Wissenschaft	Softwaredemonstrator frei bei GitHub	bis zu 3 Jahre nach Projektabschluss
Beratung von Unternehmen	Unmittelbarer Ergebnistransfer in die Industrie	Verwendung der Forschungsergebnisse in Industrieprojekten	fortlaufend (und auf Anfrage)

10 Literatur

- Baer, S.; Bakakeu, J.; Meyes, R.; Meisen, T. (2019): Multi-Agent Reinforcement Learning for Job Shop Scheduling in Flexible Manufacturing Systems. In: International Conference on Artificial Intelligence for Industries (IA4I), 2, S. 22–25.
- Baer, S.; Turner, D.; Mohanty, P.; Samsonov, V.; Bakakeu, R.; Meisen, T. (2020): Multi Agent Deep Q-Net-
work Approach for Online Job Shop Scheduling in Flexible Manufacturing. In: International Conference
on Manufacturing Systems and Multiple Machines, S. 1–9.
- Bruhn, M.; Hadwich, K. (2021): Künstliche Intelligenz im Dienstleistungsmanagement: Anwendungen, Ein-
satzbereiche und Herangehensweisen. In: Bruhn, M.; Hadwich, K. (Hrsg.): Künstliche Intelligenz im
Dienstleistungsmanagement. Band 1: Geschäftsmodelle - Serviceinnovationen - Implementierung. Wies-
baden: Springer Fachmedien, S. 1–50.
- Cadavid, J. P. U.; Lamouri, S.; Grabot, B.; Fortin, A. (2019): Machine Learning in Production Planning and
Control: A Review of Empirical Literature. In: IFAC PapersOnLine, 52-13, S. 385–390.
- Capocchi, L.; Santucci, J.-F. (2022): Discrete Event Modeling and Simulation for Reinforcement Learning Sys-
tem Design. In: Information, 13 (121), S. 1–13.
- Cook, W. J.; Cunningham, W. H.; Pulleyblank, W. R.; Schrijver, A. (1998): Combinatorial Optimization. New
York: Wiley & Sons, Inc.
- Corsten, H.; Gössinger, R. (2016): Produktionswirtschaft. Einführung in das industrielle Produktionsmanage-
ment. Berlin, Boston: Walter de Gruyter GmbH.
- Dangelmaier, W. (2009): Theorie der Produktionsplanung und -steuerung. Im Sommer keine Kirschprali-
nen? Berlin, Heidelberg: Springer-Verlag.
- François-Lavet, V.; Henderson, P.; Islam, R.; Bellemare, M. G.; Pineau, J. (2018): An Introduction to Deep
Reinforcement Learning. In: Foundations and Trends in Machine Learning, 11 (3-4).
- Gabel, T.; Riedmiller, M. (2007): Adaptive Reactive Job-Shop Scheduling with Reinforcement Learning
Agents. In: International Journal of Information Technology and Intelligent Computing, 24 (4), S. 14–18.
- Gierth, A.; Schmidt, C. (2006): Zeitdynamische Simulation in der Produktion. In: Schuh, G. (Hrsg.): Produkti-
onsplanung und -steuerung. Grundlagen, Gestaltung und Konzepte. Berlin, Heidelberg: Springer Verlag,
S. 646–681.
- Günther, H.-O.; Tempelmeier, H. (2016): Produktion und Logistik. Supply Chain und Operations Manage-
ment. Books on Demand: Norderstedt.
- Gutenschwager, K.; Rabe, M.; Spieckermann, S.; Wenzel, S. (2017): Simulation in Produktion und Logistik.
Grundlagen und Anwendungen. Berlin, Heidelberg: Springer Vieweg.
- Heidrich-Meisner, V.; Lauer, M.; Igel, C.; Riedmiller, M. (2007): Reinforcement Learning in a Nutshell. In:
Proceedings of the European Symposium on Artificial Neural Networks, 15 (1), S. 1–12.

- Heng, C.; Haifeng, X.; Smith, S. F. (2004): A reinforcement learning approach to production planning in the fabrication/fulfillment manufacturing process. In: Proceedings of the 2003 Winter Simulation Conference, 1417-1423.
- Hoschka, P. (1963): Planung des Produktionsprogramms bei konkurrierender Produktion. In: Zeitschrift für Nationalökonomie, 23 (1-2), S. 18–70.
- Jacob, H. (1996): Produktions- und Absatzprogrammplanung. In: Kern, W.; Schröder, H.-H.; Weber, J. (Hrsg.): Handwörterbuch der Produktionswirtschaft. Stuttgart: Schäffer-Poeschel Verlag, S. 1467–1483.
- Jung, H. (2016): Allgemeine Betriebswirtschaftslehre. Berlin, Boston: Walter de Gruyter GmbH.
- Kluge, S. (2011): Methodik zur fähigkeitsbasierten Planung modularer Montagesysteme. Dissertation, Universität Stuttgart (Hrsg.), Stuttgart.
- Kuhnle, A. (2020): Adaptive Order Dispatching based on Reinforcement Learning. Application in a Complex Job Shop in the Semiconductor Industry. Dissertation, Karlsruher Institut für Technologie (Hrsg.), Karlsruhe.
- Kuhnle, A.; Kaiser, J.-P.; Theiß, F.; Stricker, N.; Lanza, G. (2021): Designing an adaptive production control system using reinforcement learning. In: Journal of Intelligent Manufacturing, 32 (1), S. 855–876.
- Law, A. M. (2015): Simulation Modeling and Analysis. New York, NY: McGraw-Hill Inc.
- VDI, 2689:02-2019. Leitfaden für Materialflussuntersuchungen. Düsseldorf: VDI-Gesellschaft Produktion und Logistik (GPL).
- Lin, C.-C.; Yen-Ling Chih, D.-J. D.; Chiu, H.-T. (2019): Smart Manufacturing Scheduling With Edge Computing Using Multiclass Deep Q Network. In: IEEE Transactions on Industrial Informatics, 15 (7), S. 4276–4284.
- Liu, C.-L.; Chang, C.-C.; Tseng, C.-J. (2020): Actor-Critic Deep Reinforcement Learning for Solving Job Shop Scheduling Problems. In: IEEE Access, (8), S. 71752–71762.
- Lödding, H. (2016): Verfahren der Fertigungssteuerung. Grundlagen, Beschreibung, Konfiguration. Berlin, Heidelberg: Springer Vieweg.
- März, L.; Krug, W. (2011): Kopplung von Simulation und Optimierung. In: März, L.; Krug, W. (Hrsg.): Simulation und Optimierung in Produktion und Logistik. Praxisorientierter Leitfaden mit Fallbeispielen. Berlin, Heidelberg: Springer Verlag, S. 41–45.
- Mnih, V.; Kavukcuoglu, K.; Silver, D.; Graves, A.; Antonoglou, I.; Wierstra, D.; Riedmiller, M. (2013): Playing Atari with Deep Reinforcement Learning. In: arXiv > Computer Science > Machine Learning, (arXiv.org/abs/1312.5602), S. 1–9.
- Ostermeier, F. F.; Deuse, J. (2024): A review and classification of scheduling objectives in unpaced flow shops for discrete manufacturing. In: Journal of Scheduling, 27 (1), S. 29–49.
<https://doi.org/10.1007/s10951-023-00795-5>.
- Panzer, M.; Bender, B. (2022): Deep reinforcement learning in production systems: A systematic literature review. In: International Journal of Production Research, 60 (13), S. 4313–4341.
- Phansalkar, V. V.; Thathachar, M. A. L. (1995): Local and Global Optimization Algorithms for Generalized Learning Automata. In: Neural Computation, 7 (5), 950-973.

- Polikarpov, M.; Syberg, M.; West, N.; Deuse, J. (2025): A Simulation Framework to Study Partial Observability in Production Planning and Control with Reinforcement Learning. In: Simulation in Produktion und Logistik 2025 21. ASIM-Fachtagung (21. ASIM-Fachtagung Simulation in Produktion und Logistik), 24. bis 26. September 2025. <https://doi.org/10.25368/2025.233>.
- Rabe, M. (2000): Future of Simulation in Production and Logistics: Facts and Vision. In: Mertins, K.; Rabe, M. (Hrsg.): Simulation in Produktion und Logistik. Tagungsband der 9. ASIM-Fachtagung. Göttingen: Fraunhofer IRB Verlag, S. 21–43.
- Rauh, E. (1998): Methodische Einbindung der Simulation in die betrieblichen Planungs- und Entscheidungsabläufe. Dissertation. Bamberg: Meisenbach Verlag.
- Riedmiller, S.; Riedmiller, M. (1999): A Neural Reinforcement Learning Approach to Learn Local Dispatching Policies in Production Scheduling. In: Proceedings of the 16th International Joint Conference on Artificial Intelligence, (2), S. 764–769.
- Rinciog, A. (2023): Demystifying Reinforcement Learning Approaches for Production Scheduling. Dissertation, TU Dortmund University (Hrsg.), Dortmund.
- Schlüter, F. (2001): Hybride PPS für heterogene Fertigungsstrukturen. Wiesbaden: Gabler Verlag.
- Schuh, G.; Brandenburg, U.; Cuber, S. (2012): 2.2 Aufgaben. In: Schuh, G.; Stich, V. (Hrsg.): Produktionsplanung und -steuerung. Grundlagen der PPS. Berlin: Springer Vieweg, S. 29–81.
- Seyyedabbasi, A.; Aliyev, R.; Kiani, F.; Gulle, M. U.; Basyildiz, H.; Shah, M. A. (2021): Hybrid algorithms based on combining reinforcement learning and metaheuristic methods to solve global optimization problems. In: Knowledge-Based Systems, 223 (8), Artikel-Nr. 107044.
- Silver, D.; Hubert, T.; Schrittwieser, J.; Antonoglou, I.; Lai, M.; Guez, A.; Lanctot, M.; Sifre, L.; Kumaran, D.; Graepel, T.; Lillicrap, T.; Simonvan, K.; Hassabis, D. (2018): A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. In: Science, 362, S. 1140–1144.
- VDI, 3633:05-2018. Simulation von Logistik-, Materialfluss- und Produktionssystemen. Düsseldorf: VDI-Gesellschaft Produktion und Logistik (GPL).
- Sutton, R. S.; Barto, A. G. (2018): Reinforcement Learning. An Introduction. Cambridge, London: MIT Press.
- Tassel, P.; Gebser, M.; Schekotihin, K. (2021): A reinforcement learning environment for job-shop scheduling. In: arXiv > Computer Science > Machine Learning, (arXiv:2104.03760v1), S. 1–9.
- Vahrenkamp, R.; Siepermann, C. (2008): Produktionsmanagement. München: Oldenbourg Wissenschaftsverlag GmbH.
- Vinyals, O.; Babuschkin, I.; Czarnecki, W. M.; Mathieu, M.; Dudzik, A.; Chung, J.; Choi, D. H.; Powell, R.; Ewalds, T.; Georgiev, P.; Oh, J.; Horgan, D.; Kroiss, M.; Danihelka, I.; Huang, A.; Sifre, L.; Chal, T.; Agaplou, J. P.; Jaderberg, M.; Vezhnevets, A. S.; Leblond, Rémi; Pohlen, T.; Dalibard, V.; Budden, D.; Sulsky, Y.; Molloy, J.; Paine, T. L.; Gulcehre, C.; Wang, Z.; Pfaff, T.; Wu, Y.; Ring, R.; Yogatama, D.; Wünsch, D.; McKinney, K.; Smith, O.; Schaul, T.; Lillicrap, T.; Kavukcuoglu, K.; Hassabis, D.; Apps, C.; Silver, D. (2019): Grandmaster level in StarCraft II using multi-agent reinforcement learning. In: Nature, 575, 350-369.
- Walter, M.; Hahn, A. (2015): Unterstützung von datengetriebenen Prozessschritten in Simulationsstudien durch Verwendung multidimensionaler Datenmodelle. In: Ritter, N.; Henrich, A.; Lehner, W.; Thor, A.;

- Friedrich, S.; Wingerath, W. (Hrsg.): Datenbanksysteme für Business, Technologie und Web (BTW). Workshopband. Bonn: Gesellschaft für Informatik e.V., S. 243–252.
- Waschneck, B.; Reichstaller, A.; Belzner, L.; Altenmüller, T.; Bauernhansl, T.; Knapp, A.; Kyek, A. (2018): Optimization of global production scheduling with deep reinforcement learning. In: *Procedia CIRP*, 72 (1), S. 1264–1269.
- Wenzel, H.; Smit, D.; Sardesai, S. (2019): A literature review on machine learning in supply chain management. In: Kersten, W.; Blecker, T.; Ringle, C. (Hrsg.): *Artificial Intelligence and Digital Transformation in Supply Chain Management. Innovative Approaches for Supply Chains*. Berlin: epubli GmbH, S. 413–441.
- Wenzel, S.; Collisi-Böhmer, S.; Pitsch, H.; Rose, O.; Weiß, M. (2008): *Qualitätskriterien für die Simulation in Produktion und Logistik. Planung und Durchführung von Simulationsstudien*. Berlin, Heidelberg: Springer-Verlag.
- West, N.; Hoffmann, F.; Schulte, L.; Moreno, V. H.; Deuse, J. (2021): An Approach for Deep Reinforcement Learning for Production Program Planning in Value Streams. In: Franke, J.; Schuderer, P. (Hrsg.): *Simulation in Produktion und Logistik. Tagungsband der 19. ASIM-Fachtagung*. Göttingen: Cuvillier Verlag, S. 345–356.
- Wiendahl, H.-P. (2014): *Betriebsorganisation für Ingenieure*. München: Carl Hanser Verlag.
- Zhang, W.; Dietterich, T. G. (1995): A Reinforcement Learning Approach to Job-Shop Scheduling. In: *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, 2, S. 1114–1120.