

**UNIVERSIDADE FEDERAL DE ITAJUBÁ
PROGRAMA DE PÓS GRADUAÇÃO EM
ENGENHARIA DE PRODUÇÃO**

**MONITORAMENTO DE MODELOS DE SIMULAÇÃO
UTILIZADOS COMO GÊMEOS DIGITAIS: UMA
ABORDAGEM BASEADA EM APRENDIZADO DE MÁQUINA
E CARTA DE CONTROLE**

Carlos Henrique dos Santos

Itajubá, Junho de 2023

**UNIVERSIDADE FEDERAL DE ITAJUBÁ
PROGRAMA DE PÓS GRADUAÇÃO EM
ENGENHARIA DE PRODUÇÃO**

Carlos Henrique dos Santos

**MONITORAMENTO DE MODELOS DE SIMULAÇÃO
UTILIZADOS COMO GÊMEOS DIGITAIS: UMA
ABORDAGEM BASEADA EM APRENDIZADO DE MÁQUINA
E CARTA DE CONTROLE**

Tese submetida ao programa de Pós-Graduação em Engenharia de Produção como parte dos requisitos para obtenção do Título de Doutor em Ciências em Engenharia de Produção.

Área: Engenharia de Produção

Orientador: Prof. Dr. José Arnaldo Barra Montevechi

Coorientador: Prof. Dr. José Antonio de Queiroz

Junho de 2023

Itajubá

**UNIVERSIDADE FEDERAL DE ITAJUBÁ
PROGRAMA DE PÓS GRADUAÇÃO EM
ENGENHARIA DE PRODUÇÃO**

Carlos Henrique dos Santos

**MONITORAMENTO DE MODELOS DE SIMULAÇÃO
UTILIZADOS COMO GÊMEOS DIGITAIS: UMA
ABORDAGEM BASEADA EM APRENDIZADO DE MÁQUINA
E CARTA DE CONTROLE**

Tese submetida à banca examinadora em 27 de Junho de 2023.

Banca examinadora:

Prof. Dr. Anderson Rodrigo de Queiroz (NCSU)

Prof. Dr. Eduardo Gomes Salgado (UNIFAL)

Prof. Dr. Antonio Fernando Branco Costa (UNIFEI)

Prof. Dr. Rafael de Carvalho Miranda (UNIFEI)

Prof. Dr. José Arnaldo Barra Montevechi (orientador)

Prof. Dr. José Antonio de Queiroz (coorientador)

Itajubá

2023

DEDICATÓRIA

De maneira especial, à minha namorada Alice, aos meus pais, Maria e Raimundo e à minha família.

AGRADECIMENTOS

Hoje sou grato, primeiramente, a Deus pelo dom da vida e por iluminar meus caminhos com pessoas maravilhosas, oportunidades valiosas e experiências inesquecíveis. Grato também pelo privilégio de me tornar doutor e pela responsabilidade de disseminar a paixão que tenho pela educação, pelo ensino e por ser professor.

Sou grato aos meus amados pais, Maria e Raimundo, por todo amor, cuidado, apoio e todos os mais incontáveis valores e ensinamentos desta vida. Por me permitirem sonhar tão alto e, mais ainda, por não pouparem esforços para me proporcionar a base mais sólida e linda.

Grato à minha querida namorada, companheira e futura esposa Alice, por aceitar o desafio de sonharmos juntos. Grato pelo amor, pela confiança e, acima de tudo, por escrever comigo esta e as próximas páginas dessa vida.

Grato aos meus queridos irmãos, Eduardo, Luciano e Eliana, por todo apoio nestes longos anos de caminhada, pela paciência, suporte e pela amizade diária de vocês, sem a qual eu jamais chegaria até aqui.

Grato ao meu orientador, Professor José Arnaldo, pela confiança, amizade, oportunidades, conselhos, orientações e pelo exemplo. Sua entrega como professor é inspiradora e me faz ter certeza de que escolhi a profissão certa. Mais que um orientador, ganhei um “paizão”.

Grato ao meu coorientador, Professor Queiroz, por despertar em mim o interesse pela carreira acadêmica e por toda confiança, amizade, conselhos e oportunidades. Desde o TCC até a tese, seu nome está gravado na minha formação como pessoa e profissional.

Grato também ao professor Miranda. Sua porta esteve sempre aberta para conselhos, dicas, orientações e, além de todos os trabalhos e histórias, levarei também sua amizade.

Grato aos meus amigos e demais professores do grupo de pesquisa NEAAD, do Programa de Pós-Graduação em Engenharia de Produção e também da graduação, os quais fizeram parte desta etapa. Em especial, grato aos meus companheiros de laboratório, prof. André, Serafim, João, Wesley, Alyson, Guilherme, Milena, André, Gustavo e Afonso. Graças à companhia de vocês, a caminhada até aqui foi mais prazerosa e vitoriosa.

Grato à querida UNIFEI, por me permitir uma formação que vai muito além do âmbito profissional. Ao INATEL, onde criei e crio cada vez mais amor pela profissão de professor.

Por fim, agradeço ao CNPq e à Advantech, objeto de estudo deste trabalho, por todo o incentivo dado a esta pesquisa.

A todos, muito obrigado!

EPÍGRAFE

“Feliz aquele que transfere o que sabe e aprende o que ensina.”

Cora Coralina

RESUMO

O uso de modelos de simulação como Gêmeos Digitais (GD) vem ganhando destaque nos últimos anos e representa uma revolução na tomada de decisão em processos produtivos, sendo uma solução chave no contexto da chamada Indústria 4.0. Neste sentido, nota-se decisões cada vez mais rápidas e eficientes a partir do espelhamento do comportamento dos sistemas físicos por meio de sensores, equipamentos inteligentes, sistemas de gestão e bases de dados. Os modelos utilizados como GD são atualizados periodicamente, em tempo real ou quase real e de acordo com as mudanças físicas, e fornecem orientações ou comandos para a tomada de decisões. Por outro lado, apesar da grande aplicabilidade dessa abordagem, destacam-se desafios relacionados à validade dos modelos de simulação ao longo do tempo, uma vez que as abordagens tradicionais de validação não consideram a atualização periódica do modelo. Garantir a validade dos GDs é vital, uma vez que tal utilização costuma envolver decisões de grande impacto para os sistemas produtivos em geral. Além disso, embora seja um campo de pesquisa com grande importância tanto para pesquisadores quanto para profissionais da área, nota-se que há ainda uma lacuna quanto a métodos visando o monitoramento da validade dos GDs. Dessa forma, de modo a contribuir com a literatura e preencher tal lacuna existente, o presente trabalho propõe uma abordagem baseada na avaliação periódica dos modelos de simulação utilizados como GDs por meio de Aprendizado de Máquina e carta de controle. Sugere-se uma ferramenta de monitoramento baseada no classificador *K-Nearest Neighbors* (K-NN), combinado com a carta de controle p , visando avaliar periodicamente a validade dos modelos de GDs. Inicialmente, a abordagem proposta foi testada em diversos casos teóricos de forma a se avaliar o funcionamento da ferramenta em situações onde o ambiente físico difere significativamente do virtual, fato que representaria um possível caso onde o GD não é válido. Neste caso, foram emulados dados correspondentes aos ambientes físico e virtual considerando distribuições de probabilidade padronizadas. Além disso, a ferramenta também foi implementada em dois objetos de estudo reais, atuando como um suplemento visando tornar os GDs mais robustos e confiáveis. Neste caso, foram adotados GDs já implementados e em fase operacional. O primeiro objeto de estudo refere-se a um modelo que apoia decisões de planejamento operacional em uma empresa de médio porte do ramo de confecções, cujos processos são majoritariamente manuais. Já o segundo objeto de estudo refere-se a um GD implementado em uma célula de produção automatizada e que opera próximo de tempo real, permitindo a avaliação dos principais parâmetros do processo. A ferramenta se mostrou capaz de monitorar o funcionamento de ambos os GDs e identificar possíveis causas especiais que podem comprometer seus resultados e, conseqüentemente, sua validade. Por fim, destaca-se a ampla aplicabilidade da ferramenta, que pode ser utilizada em diferentes abordagens de GD, incluindo modelos de simulação com diferentes características de conexão, integração e complexidade. Neste caso, a abordagem proposta opera independentemente das características dos GDs, ou seja, contempla modelos que operam em tempo real ou quase real, considerando sistemas físicos automatizados ou manuais e abrange sistemas com diferentes níveis de complexidade.

Palavras-chave: Simulação, Gêmeo Digital, validade do modelo, Aprendizado de Máquina, Cartas de Controle.

ABSTRACT

The use of simulation models as Digital Twins (DTs) has been standing out in recent years and represents a revolution in decision-making in production processes, being a key solution in the context of the so-called Industry 4.0. In this sense, we highlight increasingly faster and more efficient decisions from the mirroring of the behavior of physical systems through sensors, intelligent equipment, management systems and databases. The models used as DTs are updated periodically, in real or near real time according to physical changes, and provide guidelines or commands for decision making. On the other hand, despite the great applicability of this approach, challenges related to the validity of simulation models over time stand out, since traditional validation approaches do not consider the periodic update of the model. Ensuring the validity of DTs is essential, since it usually involves decisions of great impact for production systems. In addition, although it is a field of research with great importance for both researchers and professionals, we noted that there is still a gap in terms of methods aimed at monitoring the validity of DTs. Therefore, in order to contribute to the literature and fill this gap, the present work proposes an approach based on the periodic evaluation of simulation models used as DTs through Machine Learning and control chart. We suggest a monitoring tool based on the K-Nearest Neighbors (K-NN) classifier, combined with the p control chart, in order to periodically assess the validity of DT models. Initially, the proposed approach was tested in several theoretical cases in order to evaluate the functioning of the tool in situations where the physical environment differs significantly from the virtual one, a fact that would represent a possible case where the DT is not valid. In this case, data corresponding to the physical and digital environments were emulated considering standardized probability distributions. Furthermore, the tool was also implemented in two real objects of study, acting as a supplement to make DTs more robust and reliable. In this case, DTs already implemented and in the operational phase were adopted. The first object of study refers to a model that supports operational planning decisions in a medium-sized company of a clothing industry, whose processes are mostly manual. The second object of study refers to a DT implemented in an automated production cell that operates in near real time, allowing the evaluation of the main process parameters. The tool proved to be capable of monitoring the functioning of both DTs and identifying possible special causes that could compromise its results and, consequently, its validity. Finally, the broad applicability of the tool is highlighted, which can be used in different approaches of DT, including simulation models with different characteristics of connection, integration, and complexity. In this case, the proposed approach operates independently of the characteristics of the DTs, including models that operate in real or near real time, considering automated or manual physical systems and covers systems with different levels of complexity.

Keywords: *Simulation, Digital Twin, model validity, Machine Learning, Control Chart.*

LISTA DE FIGURAS

Figura 1.1– Arquitetura geral quanto ao uso da simulação como Gêmeo Digital	18
Figura 2.1– Abordagens quanto à coleta e processamento de dados na simulação	29
Figura 2.2 – Estrutura geral do Gêmeo Digital para apoio à decisão	38
Figura 2.3 – Estrutura geral das técnicas de Aprendizado de Máquina	43
Figura 2.4 – Influência do valor de ‘K’ na classificação pelo K-NN	45
Figura 2.5 – Estrutura geral da carta de controle	47
Figura 3.1– Classificação da pesquisa	53
Figura 3.2 – Etapas seguidas na construção dos modelos de simulação	54
Figura 3.3 – Etapas seguidas na execução dos experimentos	55
Figura 4.1 – Lacunas associadas ao uso da simulação como Gêmeo Digital	58
Figura 4.2 – Procedimentos realizados nas etapas de planejamento e pesquisa/triagem	60
Figura 4.3 – Publicações quanto à simulação como Gêmeo Digital	61
Figura 4.4 – Publicações quanto à simulação como Gêmeo Digital (estratificado)	62
Figura 4.5 – Principais áreas de atuação quanto à simulação como Gêmeo Digital	63
Figura 4.6 – Principais objetivos quanto à simulação como Gêmeo Digital	64
Figura 4.7 – Conexão entre os modelos de simulação e os sistemas físicos	66
Figura 4.8 – Atividades desempenhadas pela ferramenta de monitoramento	73
Figura 4.9 – Ciclo de uso da ferramenta de monitoramento	77
Figura 4.10 – Comparação dos dados físicos e virtuais (Objeto de estudo I)	80
Figura 4.11 – Comparação dos dados físicos e virtuais (Objeto de estudo IV)	81

Figura 4.12 – Procedimento adotado para obtenção do valor ótimo de ‘K’	82
Figura 4.13 – Carta de controle para o objeto de estudo I	83
Figura 4.14 – Carta de controle para o objeto de estudo II	83
Figura 4.15 – Carta de controle para o objeto de estudo III	83
Figura 4.16 – Carta de controle para o objeto de estudo IV	84
Figura 4.17 – Carta de controle para o objeto de estudo V	84
Figura 4.18 – Replicação para identificar a causa especial (objeto de estudo I)	85
Figura 4.19 – Planta baixa e fluxo da linha utilizada como objeto de estudo VI	86
Figura 4.20 – Modelo de SED do objeto de estudo VI	86
Figura 4.21 – Estrutura de funcionamento do Gêmeo Digital do objeto de estudo VI ..	87
Figura 4.22 – Carta de controle para o objeto de estudo VI	89
Figura 4.23 – Planta baixa e fluxo da linha utilizada como objeto de estudo VII	90
Figura 4.24 – Modelo de SED do objeto de estudo VII	90
Figura 4.25 – Estrutura de funcionamento do Gêmeo Digital do objeto de estudo VII .	91
Figura 4.26 – Carta de controle para o objeto de estudo VII	92
Figura 4.27 – Avaliação dos objetos de estudo VI e VII por testes de hipóteses	95
Figura 4.28 – Mapa de correlação entre as variáveis do objeto de estudo VI	96

LISTA DE QUADROS

Quadro 2.1 – Simulação em aplicações tradicionais vs. Gêmeo Digital	39
Quadro 3.1 – Etapas consideradas para o desenvolvimento da pesquisa	56
Quadro 4.1 – Vantagens, desafios e oportunidades da simulação como Gêmeo Digital	69

LISTA DE TABELAS

Tabela 4.1 – Relação entre a acurácia do K-NN e a validade do Gêmeo Digital	74
Tabela 4.2 – Parâmetros das distribuições adotadas nos objetos de estudo teóricos	79

LISTA DE ABREVIÇÕES

AM – Aprendizado de Máquina

AGV – *Automated Guided Vehicle* (Veículo Guiado por Automação)

ARL – *Average Run Length* (Número médio de amostras até o sinal)

CNC – Comando Numérico Computadorizado

CPS – *Cyber-physical Systems* (Sistemas Ciber-físicos)

ERP – *Enterprise Resource Planning* (Sistema de gestão empresarial)

GD – Gêmeo Digital

GPS – *Global Positioning System* (Sistema de Posicionamento Global)

IA – Inteligência Artificial

IOT – *Internet of Things* (Internet das Coisas)

K-NN – *K-Nearest Neighbors*

ML – *Machine Learning* (Aprendizado de Máquina)

NMA – Número Médio de Amostras até o Sinal

QP – Questão de Pesquisa

RL – Revisão da Literatura

RNA – Redes Neurais Artificiais

RSL – Revisão Sistemática da Literatura

SBA – Simulação Baseada em Agentes

SED – Simulação a Eventos Discretos

SD – Sistema de Dados

SF – Sistema Físico

SS – Sistema de Serviço

SV – Sistema Virtual

TI – Tecnologia de Informação

SUMÁRIO

1.INTRODUÇÃO	16
1.1. Contextualização.....	16
1.2. Justificativa	18
1.3. Problema de pesquisa e objetivos	21
1.4. Condições de contorno da pesquisa	21
1.5. Estrutura do trabalho.....	22
2. FUNDAMENTAÇÃO TEÓRICA.....	23
2.1. Simulação para apoio à tomada de decisões	24
2.1.1. Conceitos e definições gerais	24
2.1.2. Simulação a Eventos Discretos para apoio a decisões	26
2.1.3. Da abordagem tradicional à indústria moderna.....	28
2.2. Indústria 4.0 e o papel dos Gêmeos Digitais.....	31
2.2.1. Conceitos e definições	31
2.2.2. Pilares da Indústria 4.0 e o papel dos Gêmeos Digitais	33
2.3. Uso da simulação como Gêmeo Digital de processos	38
2.3.1. Conceitos e definições	38
2.3.2. Validade dos modelos de simulação utilizados como Gêmeos Digitais	40
2.4. Aprendizado de Máquina e o <i>K-Nearest Neighbors</i> (K-NN).....	42
2.4.1. Conceitos e Definições	42
2.4.2. K-NN como classificador	44
2.5. Cartas de controle como mecanismo de monitoramento	46
2.5.1. Conceitos e Definições	46
2.5.2. Carta de controle <i>p</i>	48
3. MÉTODO DE PESQUISA	50
3.1. Classificação da pesquisa.....	50
3.1.1. Quanto à natureza	51
3.1.2. Quanto ao objetivo	51
3.1.3. Quanto à abordagem.....	51
3.1.4. Quanto ao método.....	52
3.2. Método	53
3.2.1. Modelagem e Simulação	53
3.2.2. Experimentação	54

3.3. Etapas da pesquisa	55
4. APLICAÇÃO.....	56
4.1. Estado da arte acerca do uso da simulação como Gêmeo Digital: uma Revisão Sistemática da Literatura	56
4.1.1. Planejamento	57
4.1.2. Pesquisa/triagem.....	59
4.1.3. Análise e apresentação.....	60
4.2. Proposta de monitoramento dos modelos de simulação utilizados como Gêmeos Digitais.....	71
4.2.1. Definição das variáveis de avaliação.....	71
4.2.2. Criação e configuração da interface de monitoramento	72
4.2.3. Monitoramento periódico	76
4.3. Resultados Experimentais.....	78
4.3.1. Objetos de estudo teóricos (Objetos I - V)	78
4.3.2. Objetos de estudo reais (Objetos VI e VII)	85
5. CONCLUSÕES.....	96
5.1. Síntese dos Resultados	97
5.2. Limitações da pesquisa	99
5.3. Recomendações para trabalhos futuros.....	100
APÊNDICE A - Publicações	101
i. Periódicos	101
ii. Livros.....	103
iii. Congressos.....	103
REFERÊNCIAS	105

1. INTRODUÇÃO

1.1. Contextualização

A simulação computacional se consolidou nas últimas décadas como uma valiosa técnica de apoio à decisão em processos produtivos (MOURTZIS, 2020; SCHEIDEGGER *et al.*, 2018). Tal relevância se deve à sua capacidade de análise de sistemas complexos aliada à flexibilidade e ao baixo custo, uma vez que a simulação permite avaliações de sistemas sem a necessidade de interferência direta nos mesmos (BANKS *et al.*, 2010; GREASLEY e OWEN, 2018). Por meio de modelos virtuais é possível representar o comportamento de sistemas físicos com suas principais características, permitindo a realização de experimentos, a comparação de cenários e a otimização de decisões. Neste caso, destacam-se aplicações em diversos setores, incluindo logística, saúde, militar e, principalmente, manufatura (MOURTZIS, 2020; NEGAHBAN e SMITH, 2014). Além disso, destacam-se como sendo os principais tipos de simulação adotados em sistemas produtivos a Simulação a Eventos Discretos (SED) e a Simulação Baseada em Agentes (SBA) (SCHEIDEGGER *et al.*, 2018).

Entretanto, apesar de sua grande aplicabilidade, nota-se que a simulação, considerando seus principais tipos, vem passando por diversas mudanças nos últimos anos (RODIČ, 2017). Barlas e Heavey (2016) e Rodič (2017) destacam que o desenvolvimento de soluções tecnológicas como sensores e dispositivos inteligentes, a alta capacidade de processamento computacional e a integração facilitada entre os ambientes físicos e virtuais impulsionaram uma grande mudança com relação ao uso da simulação. Neste contexto, nota-se que abordagens tradicionais de simulação, marcadas por modelos complexos e análises pontuais, vem perdendo espaço para modelos integrados e sincronizados com os sistemas físicos, com interface amigável para tomada de decisões e de uso contínuo (GOODALL, SHARPE e WEST, 2019; MONTEVECHI *et al.*, 2020). Dessa forma, pode-se dizer que a simulação vive a era do chamado Gêmeo Digital (TIACCI, 2020; ZHUANG, LIU e XIONG, 2018)

O termo Gêmeo Digital (GD), do inglês *Digital Twin*, refere-se a cópias virtuais capazes de se conectar a sistemas físicos, espelhando-os a partir de seus dados e informações e orientando a tomada de decisão (WRIGHT e DAVIDSON, 2020). Tal conceito foi proposto por Shafto *et al.* (2010) e se referia, inicialmente, a cópias virtuais inteligentes dos equipamentos pertencentes à Agência Aeroespacial dos Estados Unidos, NASA. Desde sua criação, nota-se que os GDs vêm sendo amplamente utilizados nas mais diversas áreas, tais como manufatura, saúde e serviços e contemplam cópias virtuais tanto de produtos como de processos (WRIGHT e DAVIDSON,

2020). Além disso, grande parte da ascensão do uso de GDs se deve ao fato deste conceito estar atrelado ao que seria a quarta revolução industrial, a chamada Indústria 4.0 (SEMERARO *et al.*, 2021). Neste caso, os modelos virtuais propiciam a tomada de decisão baseada na digitalização de processos, um dos pilares dessa nova era industrial juntamente com outras soluções e desenvolvimentos tecnológicos (XU, XU e LI, 2018; ZHONG *et al.*, 2017).

Tao e Zhang (2017) destacam que a adoção de GDs pelos tomadores de decisão é uma tendência inevitável e, neste caso, Terkaj *et al.* (2019) revelam que a simulação se apresenta como uma alternativa mais barata e flexível na concepção dos GDs de processos se comparada às soluções comerciais disponíveis. Dessa forma, ao conectar o modelo de simulação com sistemas físicos, têm-se uma cópia virtual altamente sincronizada e sensível a mudanças reais, tornando a tomada de decisões alinhada ao estado atual dos processos (BEREGI, SZALLER e KÁDÁR, 2018). Além disso, Santos *et al.* (2022a) defendem que o uso da simulação como GD potencializa o poder da ferramenta, tornando-a capaz de reconhecer as mudanças físicas dos sistemas no decorrer do tempo e, conseqüentemente, apoiar de maneira mais eficiente a tomada de decisões. Neste caso, além da sua integração com os sistemas físicos, destacam-se modelos com baixo nível de abstração, alta capacidade gráfica, de fácil manipulação e integrados com outras ferramentas e técnicas de decisão (MOURTZIS, 2020; RODIČ, 2017; SORMAZ e MALIK, 2018).

Diversos trabalhos na literatura abordam o uso da simulação como GD, com destaque para os últimos anos (BEREGI, SZALLER e KÁDÁR, 2018; GRUBE, MALIK e BILBERG, 2019; KARAKRA *et al.*, 2018; MURPHY *et al.*, 2020; PRAJAPAT *et al.*, 2019; SANTOS *et al.*, 2022a; STERINGER *et al.*, 2019; TERKAJ *et al.*, 2019; VIJAYAKUMAR *et al.*, 2019). Neste caso, observa-se diferentes abordagens quanto à construção de GDs por meio da Simulação. Nota-se a adoção tanto de GDs autônomos, capazes de executar comandos nos sistemas físicos, como não autônomos, os quais fornecem diretivas para a tomada de decisões (DONHAUSER *et al.*, 2018; MURPHY *et al.*, 2020). Além disso, observa-se GDs que operam em tempo real, ou seja, que se atualizam instantaneamente de acordo com as mudanças nos sistemas físicos, bem como abordagens em tempo quase real, as quais baseiam-se em modelos que se atualizam em intervalos de tempo definidos de acordo com a frequência da tomada de decisões (CASSETTARI *et al.*, 2017; WAN *et al.*, 2014). A Figura 1.1 esquematiza o uso da simulação como Gêmeo Digital.

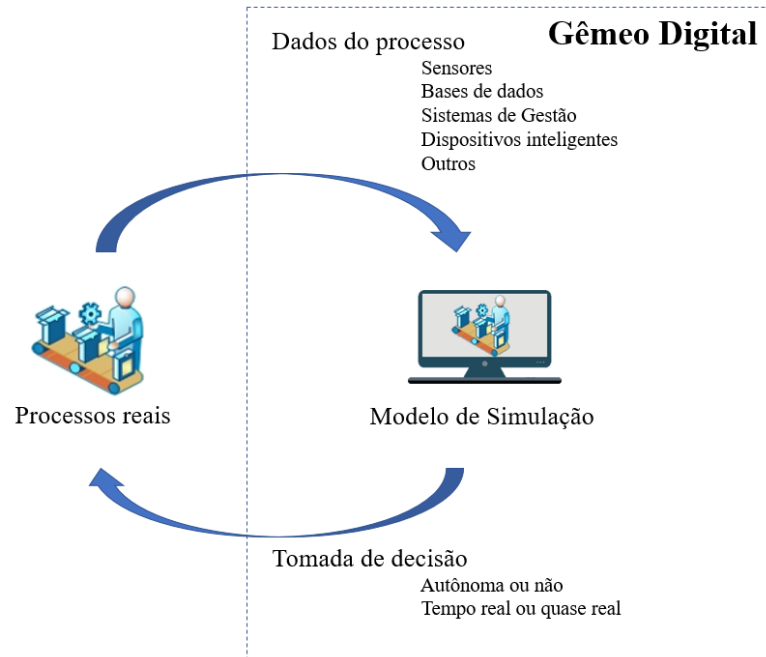


Figura 1.1 – Arquitetura geral quanto ao uso da simulação como Gêmeo Digital

Fonte: elaborado pelo próprio autor

Embora a adoção de GDs de processos seja uma solução promissora quanto à tomada de decisões e amplamente adotada por pesquisadores e praticantes, há ainda diversas questões pouco exploradas na literatura no que se refere às características dos modelos. Neste caso, destaca-se a necessidade por trabalhos que abordem métodos para a construção dos modelos virtuais, a integração e conexão entre os modelos virtuais e os sistemas físicos, a segurança associada ao uso dos GDs para tomada de decisão, o nível de inteligência dos modelos virtuais e sua utilização integrada com técnicas e ferramentas auxiliares para tomada de decisões mais eficientes, a necessidade de se garantir a validade dos modelos durante sua operação, dentre outros (EYRE, SCOTT e FREEMAN, 2018; GOODALL, SHARPE e WEST, 2019; SANTOS *et al.*, 2022b; TAO e ZHANG, 2017; WRIGHT e DAVIDSON, 2020).

1.2. Justificativa

A adoção de GDs como mecanismo de apoio a decisões em processos produtivos ainda representa um campo fértil de pesquisa (MOURTZIS, 2020; SANTOS *et al.*, 2022a; TAO e ZHANG, 2017). Dentre as diversas oportunidades de pesquisa destacadas, nota-se uma questão crítica quanto ao uso dos GDs: a validade dos modelos. Neste caso, garantir a validade dos modelos virtuais no decorrer do tempo é fundamental, uma vez que o uso de GDs está geralmente associado a decisões de alto impacto nos processos produtivos (WRIGHT e

DAVIDSON, 2020). Logo, ao se considerar o caráter dinâmico dos processos físicos e a atualização periódica do ambiente virtual, são necessários métodos e técnicas visando garantir a validade dos modelos durante sua operação como GD (SANTOS *et al.*, 2022b; TAO e ZHANG, 2017; WRIGHT e DAVIDSON, 2020).

Considerando abordagens tradicionais de simulação, nota-se que métodos visando garantir a validade dos modelos durante sua etapa de construção são amplamente difundidos na literatura (BANKS *et al.*, 2010; SARGENT, 2013). Por outro lado, ao considerar seu uso como GDs, garantir a validade dos modelos de simulação durante sua operação é igualmente importante e crucial, além de representar um campo promissor para pesquisas (HARPER e MUSTAFEE, 2019; ONGGO *et al.*, 2018; SANTOS *et al.*, 2022b). Neste caso, Meng *et al.*, (2013), afirmam que garantir a validade de modelos que se modificam no decorrer do tempo é ainda mais desafiador e complexo. Tal fato se deve a alguns fatores, tais como a dificuldade associada à periodicidade ideal de avaliação do modelo, a definição da ferramenta adequada para tal, bem como a complexidade do modelo e dos sistemas físicos que compõem o GD, podendo resultar em diversas variáveis para se avaliar simultaneamente.

Segundo Tao e Zhang (2017), a base para se garantir a validade dos GDs durante sua operação são rotinas sistemáticas e periódicas de comparação entre dados dos sistemas físicos e os resultados dos modelos virtuais. Nesse caso, de forma a comparar dados físicos e virtuais, destaca-se a possibilidade do uso de classificadores, ou mais precisamente do chamado *K-Nearest Neighbors* (K-NN). O K-NN é uma técnica de Aprendizado de Máquina, ou do inglês *Machine Learning*, amplamente usada para classificação e previsão e é conhecido por sua simplicidade e eficácia (LEE *et al.*, 2020; WANG, TSAI e LIN, 2021). A partir de um conjunto de dados conhecidos, o algoritmo K-NN pode aprender com os padrões e prever a classificação de novos dados (KUMAR *et al.*, 2020). Neste caso, ao se comparar dados dos sistemas físicos com dados advindos dos modelos digitais, espera-se que o classificador não consiga diferenciá-los e, conseqüentemente, este terá uma acurácia baixa frente à alta correspondência do GD com os sistemas físicos.

Ao se considerar o uso do K-NN para avaliar a validade do GD, pode-se superar dificuldades relacionadas à complexidade do sistema, visto que é possível a avaliação de diversas variáveis simultaneamente e por meio de somente um indicador, a acurácia do classificador. Porém, é necessário ainda considerar a avaliação do GD no decorrer do tempo e, neste caso, sugere-se a adoção da carta de controle visando monitorar a acurácia do K-NN e, conseqüentemente, a validade do GD, durante sua operação. As cartas de controle compõem a principal técnica de monitoramento de sistemas sujeitos a variabilidade no decorrer do tempo, sendo amplamente

utilizadas nas mais diversas áreas de aplicação (ABBAS *et al.*, 2019; CHUKHROVA e JOHANNSEN, 2019). Neste caso, considerando que a acurácia do classificador é obtida a partir da proporção de acertos do K-NN, é proposta a adoção da carta de controle p (MONTGOMERY, 2009) combinada com o classificador visando complementar a avaliação do GD durante sua operação.

Além das observações realizadas anteriormente, esta tese fundamenta-se também nas seguintes justificativas:

- Ineditismo quanto ao propósito de monitoramento de modelos de simulação durante sua operação como GDs. Por meio de uma revisão sistemática da literatura considerando trabalhos publicados até meados de 2020 em revistas científicas e congressos de impacto, nota-se que nenhum trabalho propõe um método sistemático e periódico de monitoramento dos modelos. Somente um, CHO *et al.* (2019), dentre os diversos trabalhos analisados adota a avaliação periódica do modelo por meio de testes de hipótese realizados a cada tomada de decisão. Porém, trata-se de uma abordagem simplificada e limitada considerando os desafios relacionados à validade do GD, tal como já descrito anteriormente. Outros resultados da revisão sistemática desenvolvida são apresentados no início do capítulo 4 da presente tese.
- Ineditismo quanto às ferramentas e técnicas adotadas. Neste caso, o presente trabalho propõe uma ferramenta de monitoramento de modelos de simulação utilizados como GDs visando avaliar sua validade durante todo o ciclo de vida. Para tal, utiliza-se do classificador K-NN em conjunto com a carta de controle p de forma a se comparar os ambientes físicos e virtuais sistemática e periodicamente. Embora amplamente abordados na literatura no decorrer dos anos, nota-se que o K-NN e a carta de controle p são populares em abordagens individuais, sendo a integração de ambas como uma ferramenta única para monitoramento de GDs uma abordagem ainda não explorada.

Por fim, vale destacar que a abordagem de monitoramento proposta resultou em duas publicações científicas de alto impacto, fato que reforça o ineditismo do tema. Além disso, a revisão sistemática da literatura que embasou a justificativa deste tema também resultou em uma publicação científica de alto impacto e ressalta a importância do trabalho desenvolvido nesta tese. Mais detalhes sobre as publicações encontram-se ao final deste trabalho.

1.3. Problema de pesquisa e objetivos

Considerando a lacuna existente na literatura quanto a métodos e ferramentas de monitoramento dos modelos de simulação utilizados como GDs durante sua operação, bem como a importância de se garantir a validade desses modelos visando decisões mais eficientes e confiáveis, a presente tese pretende responder à seguinte pergunta de pesquisa: **é possível garantir a validade dos modelos de simulação utilizados como GD a partir de uma ferramenta de monitoramento baseada no classificador K-NN e na carta de controle p ?**

Dessa forma, o objetivo geral do presente trabalho é propor a ferramenta de monitoramento, bem como analisar a sua aplicabilidade e eficiência por meio de sua aplicação em objetos de estudo teóricos e reais. Neste caso, pretende-se avaliar as vantagens e limitações da ferramenta proposta, bem como destacar seu papel fundamental como um suplemento para os GDs de forma a torná-los mais confiáveis sob o ponto de vista da tomada de decisões. Além disso, o presente trabalho possui os seguintes objetivos específicos:

- Realizar uma Revisão Sistemática da Literatura (RSL) visando explorar o estado da arte sobre o uso da simulação como GD de processos;
- Desenvolver uma interface que servirá de base para a ferramenta de monitoramento, a qual será baseada no classificador K-NN e na carta de controle p ;
- Desenvolver/definir objetos de estudo teóricos e reais visando testar a ferramenta proposta;
- Analisar a aplicabilidade e eficiência da abordagem proposta quando submetida a situações específicas no que se refere à operação do GD;
- Validar a abordagem proposta como um suplemento para GDs construídos por meio da simulação.

1.4. Condições de contorno da pesquisa

Tendo em vista os objetivos deste trabalho, é importante ressaltar algumas considerações acerca do escopo da pesquisa, as quais servem como referência para as análises, avaliações e conclusões apresentadas nesta tese. Primeiramente, é importante frisar que, embora haja outras alternativas para a concepção dos GDs, grande parte associada a pacotes comerciais que acompanham equipamentos e máquinas, o foco deste trabalho limita-se aos GDs construídos por meio da simulação computacional. Neste caso, trata-se de uma alternativa mais barata e flexível para a digitalização dos processos, a base para a tomada de decisões baseada nos GDs (LU *et al.*, 2019; TERKAJ *et al.*, 2019). Além disso, o foco deste trabalho não está na

construção dos GDs, mas sim em sua operação para auxílio à tomada de decisões. Logo, etapas e atividades relacionadas à construção do modelo virtual por meio da simulação não serão abordadas em detalhes neste trabalho e serão baseadas em outros trabalhos científicos (MONTEVECHI *et al.*, 2020; SANTOS *et al.*, 2022a).

Outro ponto a ser considerado refere-se ao tipo de GD abordado. Neste caso, este trabalho refere-se somente aos GDs de processos (TAO e ZHANG, 2017), não englobando os modelos focados em produtos (ex. máquinas e equipamentos), embora ambas as abordagens sejam amplamente tratadas na literatura (WRIGHT e DAVIDSON, 2020). Neste trabalho é considerado processo toda atividade produtiva que possui entradas, processamento a partir da utilização de recursos e que geram saídas (produtos ou serviços). Além disso, ao se considerar o uso da simulação como GDs em processos produtivos, destacam-se a SED, a SBA e as abordagens híbridas (que integram o uso de ambas) como sendo os principais tipos de simulação adotados, com grande destaque para a SED que representa cerca de 92% das aplicações (SANTOS *et al.*, 2022b). Este e outros resultados são detalhados na revisão sistemática da literatura presente do capítulo 4 deste trabalho. Diante dessa expressiva presença da SED, o foco principal deste trabalho recai sobre a sua utilização para a concepção de GDs em processos produtivos.

Por fim, vale destacar que este trabalho se limita ao uso do classificador K-NN e da Carta de Controle p como base para o planejamento e implementação da ferramenta de monitoramento. Neste caso, embora haja outras alternativas quanto às técnicas de Aprendizado de Máquina e Cartas de Controle, este trabalho não objetiva testar e implementar a melhor ferramenta possível, fato que demandaria diversos testes visando identificar quais dessas alternativas resultam em maiores eficiências no monitoramento. Por outro lado, pretende-se preencher uma lacuna de pesquisa com base em técnicas amplamente difundidas por pesquisadores e profissionais: o classificador K-NN e a Carta p . Além disso, optou-se por sugerir tais análises em possíveis trabalhos futuros.

1.5. Estrutura do trabalho

A presente tese está estruturada em cinco capítulos. O **Capítulo 1** apresentou uma introdução do trabalho, abordando a contextualização do tema, as justificativas científicas que embasam o trabalho, o problema de pesquisa e objetivos da tese, as condições de contorno da pesquisa e, por fim, a estrutura adotada. O **Capítulo 2** apresenta a fundamentação teórica sobre os principais tópicos abordados. Primeiramente são apresentados conceitos relacionados ao uso da

simulação para apoio a decisões e sua evolução no decorrer do tempo, seguido por discussões acerca da chamada Indústria 4.0 e o papel da simulação como GD. Neste caso, foca-se em aspectos relacionados à validade do modelo de simulação quando utilizado como GD. Além disso, são apresentados conceitos e definições associadas às técnicas de Aprendizado de Máquina e ao classificador K-NN, bem como ao uso de cartas de controle como mecanismo de monitoramento.

No **Capítulo 3** é apresentado o método de pesquisa adotado, incluindo a classificação da pesquisa. Já o **Capítulo 4** apresenta, inicialmente, o estado da arte do uso da simulação como GD por meio de uma revisão sistemática da literatura. Em seguida, apresenta-se a abordagem proposta, destacando todos os detalhes da ferramenta de monitoramento de modelos de simulação como GDs. A seguir, são apresentados os resultados experimentais oriundos da aplicação da abordagem proposta em objetos de estudo teóricos e reais. Finalmente, o **Capítulo 5** contém as conclusões da pesquisa, com destaque para os principais resultados, limitações e oportunidades para trabalhos futuros. Por último, encontram-se os apêndices da tese, incluindo as publicações em periódicos e congressos resultantes deste trabalho, seguidos das referências utilizadas na pesquisa.

2. FUNDAMENTAÇÃO TEÓRICA

Neste capítulo serão apresentados, inicialmente, tópicos gerais relacionados ao uso da simulação no apoio a decisões, abordando conceitos principais, bem como suas classificações, com ênfase à SED. Além disso, serão abordados aspectos relacionados à evolução da simulação de processos no decorrer dos anos, com destaque para sua aplicação na indústria moderna. Em seguida, a chamada Indústria 4.0 e o papel da simulação como alternativa para concepção dos GDs serão abordados. Neste caso, além dos conceitos gerais e pilares principais dessa nova era industrial, são apresentados também as características principais dos chamados GDs e o papel da simulação como valiosa alternativa na sua concepção. Neste caso, são destacados conceitos relacionados à validade do modelo de simulação durante sua operação como GD. Por fim, as subseções finais apresentam aspectos relacionados às duas técnicas adotadas neste trabalho, técnicas de Aprendizado de Máquina com ênfase no classificador K-NN e o uso de cartas de controle como mecanismo de monitoramento.

2.1. Simulação para apoio à tomada de decisões

2.1.1. Conceitos e definições gerais

A simulação baseia-se na imitação de um ou mais sistemas reais por meio de modelos computacionais, a partir dos quais são formuladas inferências acerca das características representativas destes sistemas (BANKS, 1998). Law (2014) relata que a simulação é uma das principais técnicas de Pesquisa Operacional, destacando-se por suas diversas aplicações, tais como: concepção e análises de sistemas de manufatura; avaliações de sistemas financeiros; planejamento e análises de serviços; planejamento de operações logísticas e militares; dentre outras aplicações. Além disso, Banks *et al.* (2010) revelam que a construção dos modelos de simulação baseia-se em observações sobre o sistema a ser investigado seguida da representação do mesmo por meio de técnicas e relações matemáticas, lógicas e/ou simbólicas.

Com a evolução de recursos computacionais, a simulação obteve rápido crescimento e ganhou novos recursos e áreas de atuação com o passar dos anos (GOLDSMAN, NANCE e WILSON, 2010). Fishman (2001) salienta que a modelagem de processos se tornou um meio de sobrevivência para diversas áreas, como: engenharia, serviços de saúde, telecomunicações, militares, transporte, dentre outras. Para o autor, a modelagem por meio da simulação é capaz de fornecer respostas para tomada de decisões em sistemas complexos a um custo relativamente baixo, fato que contribui para a grande disseminação da simulação no decorrer dos anos. Além disso, Chwif e Medina (2015) complementam ao relatar que, apesar da complexidade dos sistemas reais devido à sua natureza dinâmica e aleatória, a simulação consegue capturar com fidelidade suas principais características e comportamentos, favorecendo a tomada de decisão mais eficiente e eficaz.

A simulação pode ser compreendida sob diferentes aspectos, a depender do critério de classificação (NANCE e SARGENT, 2002). Além disso, Law (2014) complementa que o critério de escolha quanto a abordagem a ser adotada dependerá da natureza do sistema a ser modelado, bem como dos objetivos do estudo. Considerando os objetivos da simulação, destaca-se o principal deles, a **Análise**, na qual pretende-se representar determinado sistema de forma a compreender seu comportamento e realizar inferências. Por outro lado, há também objetivos relacionados a **Treinamento**, **Avaliação de sistemas** e **Pesquisa**, os quais são relacionados ao ensino de conceitos por meio dos modelos computacionais, à avaliação de sistemas quanto ao atendimento de determinados requisitos e ao teste de comportamentos no modelo computacional de forma a se testar hipóteses, respectivamente.

Outra classificação da simulação diz respeito à sua representação de tempo e estado. Neste caso, têm-se a chamada **Simulação de Monte Carlo**, a qual baseia-se na geração de números aleatórios para simular sistemas físicos e matemáticos, sem se considerar o comportamento do sistema no decorrer do tempo. Por outro lado, ao se considerar o comportamento temporal dos sistemas, há a chamada **Simulação Contínua** e a **Simulação a Eventos Discretos**, onde são retratadas mudanças de estado dos sistemas em intervalos de tempo contínuos e discretos, respectivamente (CHWIF e MEDINA, 2015; LAW, 2014; NANCE e SARGENT, 2002). Por fim, vale ainda ressaltar a existência da chamada **Simulação Baseada em Agentes**, a qual pode ser de natureza discreta ou contínua e que se apresenta como alternativa para a modelagem de determinados sistemas. Sua diferença para os demais tipos de simulação reside nas características de modelagem, a qual é baseada em agentes que representam recursos que interagem entre si (MACAL e NORTH, 2014; SIEBERS *et al.*, 2010).

Rodič (2017) chama a atenção de que a simulação, considerando suas diferentes abordagens, consolidou-se nas últimas décadas nos mais diversos setores e áreas de atuação. Neste caso, Jahangirian *et al.* (2010) ressaltam que, parte da sua popularidade se justifica devido à capacidade de atender a diferentes camadas de decisão dentro dos sistemas empresariais. Além disso, vale destacar a possibilidade de uso de técnicas e ferramentas integradas à simulação, conhecida como abordagem híbrida, visando aprimorar seus resultados e possibilitar a tomada de decisões considerando a complexidade crescente dos sistemas físicos (NEGAHBAN e SMITH, 2014). Neste caso, destaca-se o uso da simulação integrada a técnicas de otimização, realidade virtual e aumentada, ferramentas de previsão, algoritmos de Inteligência Artificial, dentre outros (MOURTZIS, 2020). Por fim, Uriarte, Ng e Moris (2018) complementam com o fato de que a simulação evoluiu nas últimas décadas, deixando de ser uma ferramenta para especialistas para se tornar uma ferramenta focada no tomador de decisões.

Ainda quanto aos tipos de simulação, tendo em vista a tomada de decisões em processos produtivos, Scheidegger *et al.* (2018) e Uriarte, Ng e Moris (2018) destacam que a SED é o principal tipo de simulação adotado, estando presente na grande maioria de aplicações listadas na literatura. Neste caso, tal destaque se justifica devido às características principais da SED, a qual baseia-se na modelagem de sistemas que evoluem instantaneamente em pontos separados no tempo a partir de ocorrências capazes de mudar o estado desses sistemas (LAW, 2014). Além disso, Fishman (2001) reforça que a SED é capaz de fornecer importantes análises aos sistemas produtivos com erros relativamente pequenos. Ainda, considerando o uso da simulação como GDs de processos, a SED continua sendo predominante, responsável por nove a cada dez aplicações (SANTOS *et al.*, 2022b). Logo, o presente trabalho será focado

especificamente na SED e a mesma poderá ser denominada como “SED” ou simplesmente como “simulação”.

2.1.2. Simulação a Eventos Discretos para apoio a decisões

Tako e Robinson (2010) descrevem a SED de maneira simplificada como sendo a representação de entidades individuais (que podem representar pessoas, peças, equipamentos de transporte, dentre outros) que se movem através de filas, atividades e locais, em pontos discretos no tempo. Ainda, tais entidades interagem com as atividades e locais a partir do sequenciamento dos eventos programados no modelo e seguindo regras que na maioria das vezes é de natureza estocástica (OPACIC e SOWLATI, 2017). Neste caso, nota-se a grande versatilidade da técnica, abrangendo desde análises operacionais até decisões em níveis gerenciais e se destacando como uma ferramenta estratégica em processos produtivos focados tanto em produtos como em serviços (JAHANGIRIAN *et al.*, 2010; TAKO e ROBINSON, 2010). Além disso, considerando as principais etapas seguidas em projetos de SED, nota-se a presença de inúmeros métodos disponíveis na literatura (MONTEVECHI *et al.*, 2015) e Banks *et al.* (2010) apresentam as principais fases adotadas em abordagens tradicionais:

- I. Formulação do problema:** trata-se da primeira etapa adotada em projetos de simulação e o foco é garantir a clara definição dos problemas a serem abordados no projeto. Balci (2012) relata que esta etapa é imprescindível e interfere em todas as demais.
- II. Definição dos objetivos e planejamento geral do projeto:** os objetivos estão relacionados às perguntas que se deseja responder por meio da simulação, enquanto o planejamento geral do projeto foca na estimativa de recursos necessários para sua execução (BANKS *et al.*, 2010; LAW, 2009);
- III. Modelagem conceitual:** o modelo conceitual é a primeira abstração do problema e trata-se de um passo fundamental para a condução do projeto de simulação, já que será a base para a modelagem computacional (ROBINSON, 2008). Neste caso, há diversos métodos e ferramentas para representação do modelo conceitual (GABRIEL *et al.*, 2022);
- IV. Modelagem dos dados de entrada:** tal fase é constituída pela identificação, coleta e modelagem de dados necessários para o modelo de simulação. A identificação e coleta estão relacionadas ao processo que vai desde a definição da precisão, período de amostragem e formato dos dados até a coleta dos mesmos por meio de diversas técnicas

(ONGOO e HILL, 2014). Já a modelagem dos dados está relacionada à adoção de técnicas estatísticas visando representar tais dados (BILLER e GUNES, 2010);

- V. Modelagem computacional:** refere-se à etapa de tradução do modelo conceitual por meio de uma linguagem de simulação ou simulador. Nota-se que os simuladores, *softwares* específicos para simulação, dominam o mercado e há a tendência se tornarem cada vez mais customizados e flexíveis para aplicações diversas (CHWIF e MEDINA, 2015);
- VI. Verificação do modelo:** o objetivo da verificação é garantir que a programação e implementação do modelo computacional esteja correta, ou seja, de acordo com as lógicas e regras definidas no modelo conceitual (SARGENT, 2013);
- VII. Validação do modelo:** Sargent (2014) relatam que a validação do modelo visa garantir que seus resultados sejam suficientemente bons quanto à precisão e acurácia, bem como adequados para determinada aplicação. Neste caso, há diversas técnicas disponíveis na literatura, as quais abordam ferramentas qualitativas e quantitativas, bem como objetivas e subjetivas. A escolha de qual técnica utilizar irá depender das características do projeto e também da equipe envolvida no mesmo (CHWIF; MEDINA, 2015; SARGENT, GOLDSMAN e YAACOUB, 2016);
- VIII. Definição de experimentos:** Balci (2012) relata que a definição de experimentos ocorre após se obter o modelo computacional válido. Neste caso, trata-se da etapa de planejamento da execução da simulação visando obter determinadas respostas. Nesta fase são levados em conta as inferências consideradas em cada experimento, bem como as estimativas de custos;
- IX. Simulação de cenários e análise:** trata-se da execução dos experimentos previamente planejados seguida da avaliação dos indicadores de performance do sistema simulado (BANKS *et al.*, 2010);
- X. Documentação do projeto:** a documentação apropriada de um projeto de simulação é um pré-requisito essencial para a comunicação e aprendizagem no decorrer do projeto. A correta documentação viabiliza a manutenção, reutilização ou reprodução do modelo de simulação e seus resultados (OSCARSSON; MORIS, 2002; TRIEBIG; KLÜGL, 2009);
- XI. Implementação:** a fase final de projetos de simulação diz respeito à tomada de decisões baseadas nos resultados obtidos. Neste caso, Banks *et al.* (2010) relatam que o sucesso desta fase está ligado às demais etapas já descritas e trata-se da razão principal pela qual o projeto foi desenvolvido.

Ainda com relação à adoção da SED para auxílio à decisão em processos produtivos, vale destacar que a ferramenta está presente em diversas aplicações como *design* de produtos, processos e *layouts*, planejamento e controle, elaboração e análise de estratégias, alocação de recursos, treinamentos, dentre outras (JAHANGIRIAN *et al.*, 2010). Além disso, Negahban e Smith (2014) destacam a expectativa de que a aplicação da SED na concepção e operação de sistemas produtivos continue crescendo e evoluindo, tendo em vista a importância econômica do setor, bem como sua constante evolução e crescente competitividade com o passar dos anos. Os autores relatam ainda quanto à necessidade de técnicas mais eficientes para lidar com a crescente complexidade das operações produtivas, sendo de extrema importância o surgimento de abordagens diferenciadas e aprimoradas visando, sobretudo, análises e decisões mais eficientes.

2.1.3. Da abordagem tradicional à indústria moderna

Apesar da grande aplicabilidade da simulação em diversas áreas e contextos, Rodič (2017) relata que a ferramenta vem passando por mudanças significativas no decorrer dos últimos anos. O autor destaca que, diferente do que se observa em projetos tradicionais de simulação, marcados por escopos e tempos limitados, a evolução dos processos produtivos demanda decisões cada vez mais flexíveis e inteligentes e, conseqüentemente, ferramentas e técnicas que acompanham tal tendência. Neste caso, Uriarte, Ng e Moris (2018) revelam que a simulação é imprescindível ao se considerar a evolução dos sistemas produtivos e uma das tecnologias-chave neste contexto. Entretanto, apesar da relevância da simulação, Skoogh, Perera e Johansson (2012) salientam para o fato de que muitos tomadores de decisões tem falhado em usufruir das vantagens trazidas pela ferramenta. Tal fato ilustra a necessidade de mudança quanto às abordagens adotadas e, dentre os principais desafios, destaca-se o alto custo em função do tempo empregado e da necessidade de especialistas no desenvolvimento do projeto (BARLAS; HEAVEY, 2016).

Tendo em vista as dificuldades relatadas, nota-se que, no decorrer dos anos, diversas abordagens foram surgindo de forma a mitigá-las ou eliminá-las. A primeira grande mudança a ser notada refere-se à coleta e modelagem de dados de entrada para modelos de simulação. Neste caso, Barlas e Heavey (2016) relatam que, considerando abordagens tradicionais de simulação, a coleta de dados é a etapa mais demorada, podendo representar até 40% do tempo total do projeto. Além disso, tal demora se justifica devido à baixa qualidade dos dados advindos de diversas fontes, bem como à dificuldade em modelá-los. Nota-se inicialmente a tendência de se automatizar a coleta e processamento dos dados e, neste caso, pode-se

classificar tal automatização em três métodos básicos, conforme ilustra a Figura 2.1 (BARLAS e HEAVEY, 2016; SKOOGH, PERERA e JOHANSSON, 2012).

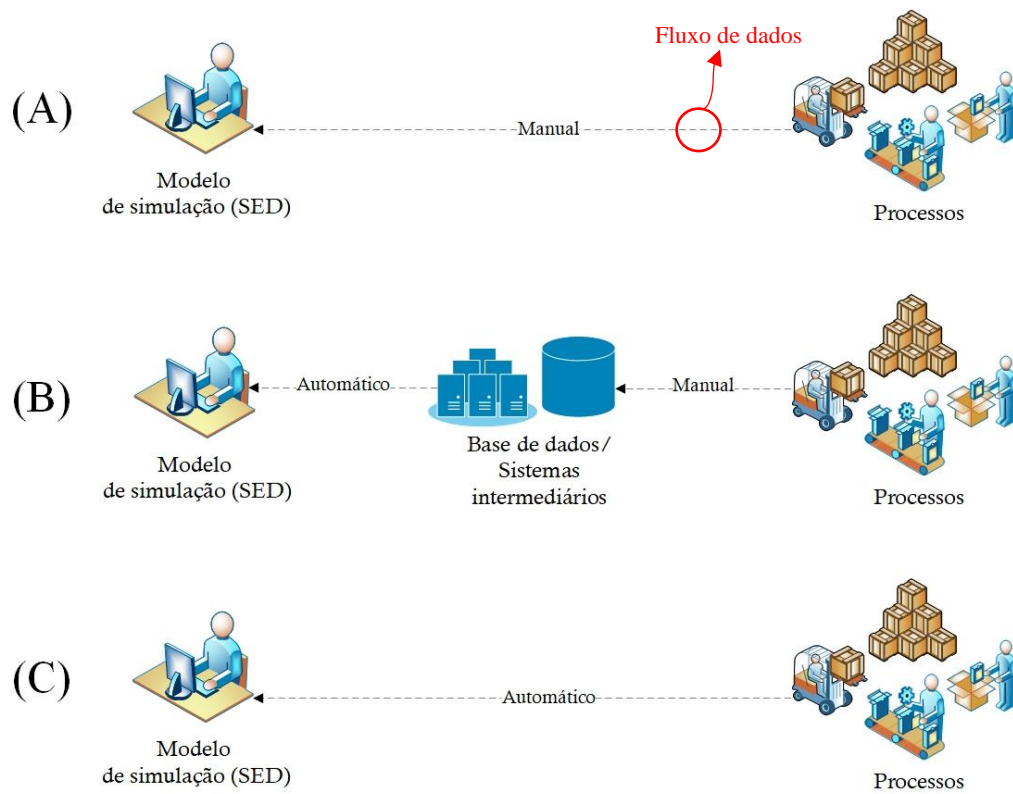


Figura 2.1 – Métodos quanto à coleta e processamento de dados na simulação
 Fonte: adaptado de Skoogh, Perera e Johansson (2012) e Barlas e Heavey (2016)

Como se observa na Figura 2.1, o método A, comum em projetos tradicionais de simulação, desprende grandes tempos para coleta e processamento de dados. Enquanto isso, há uma tendência de crescimento dos métodos B e C visando maior agilidade nos projetos de simulação (SKOOGH, PERERA e JOHANSSON, 2012). Neste caso, o que difere B e C é a presença de sistemas intermediários focados no processamento e preparação dos dados antes de serem inseridos no modelo computacional. Neste sentido, diversos autores defendem a presença desses sistemas intermediários, visto que é quase inevitável a necessidade de se trabalhar os dados de forma a torná-los válidos e adequados ao modelo (RODIČ; KANDUČ, 2015; SKOOGH; PERERA; JOHANSSON, 2012). Vale ressaltar ainda que, mesmo em casos de coleta e processamento automático de dados (método C), os modelos de simulação são atualizados perante as mudanças nos sistemas físicos, sendo, portanto, o modelo de simulação um reflexo das mudanças do sistema que ele representa (BARLAS e HEAVEY, 2016). Ainda com relação aos modelos de simulação conectados aos sistemas físicos e sensíveis às suas mudanças, nota-se que os métodos descritos na Figura 2.1 e suas variações deram origem

a diversas abordagens de simulação no decorrer dos anos. Duas abordagens bastante comuns na literatura são as chamadas “*Near Real Time Simulation*” e “*Real Time Simulation*”, ou em português, “Simulação em Tempo Quase Real” e “Simulação em Tempo Real”, respectivamente. Tratam-se de abordagens em que os modelos de simulação são abastecidos por dados advindos de processos e sistemas reais e de maneira automática, possibilitando a atualização do modelo de simulação frente às mudanças reais (VAHDATIKHAKI e HAMMAD, 2014; VAHDATIKHAKI e HAMMAD, 2015). O que difere ambas as abordagens é a frequência de atualização do modelo. Na simulação em tempo quase real, o termo “quase” implica que os modelos são atualizados em intervalos de tempo que variam de acordo com a frequência da tomada de decisão, ou seja, há um certo atraso entre as mudanças físicas e a atualização do modelo (VAHDATIKHAKI e HAMMAD, 2014). Por outro lado, na chamada Simulação em Tempo Real, tal atraso no tempo é praticamente desprezível, tornando as rodadas de atualização praticamente contínuas.

Nota-se que as abordagens de simulação em tempo quase real e tempo real vem sendo amplamente adotadas na literatura em desenvolvimentos teóricos e práticos (AKHAVIAN e BEHZADAN, 2012; CASSETTARI *et al.*, 2017; LEE, 2019; VAHDATIKHAKI e HAMMAD, 2014; ZHENG *et al.*, 2017). Saez *et al.* (2018) ressaltam que, diferentemente da abordagem tradicional de simulação, cujas análises visam a redução de custos, melhoria de produtividade e/ou análise dos sistemas, o uso da simulação sincronizada com os ambientes físicos, seja em tempo real ou quase real, visa principalmente o monitoramento e avaliação do desempenho dos sistemas durante todo seu ciclo de vida. Além disso, a escolha entre a abordagem em tempo real ou quase real depende dos objetivos do sistema e das características da decisão. Neste caso, vale considerar que abordagens em tempo real carecem de uma estrutura capaz de permitir de coleta de dados e atualização contínua do modelo, fato que pode ser um desafio tanto financeiro quanto prático para determinados casos e aplicações (SONG e ELDIN, 2012).

Outro ponto importante é que diversas novas abordagens foram surgindo a partir do conceito da simulação em tempo real ou quase real. Em primeiro lugar destaca-se a chamada “*Online Simulation*” ou “Simulação *Online*” em português. De acordo com Gupta e Sivakumar (2005) e Scholl *et al.* (2010), essa abordagem baseia-se na conexão do modelo de simulação com sistemas reais para auxiliar na tomada de decisões levando em conta o comportamento dinâmico das operações e, neste caso, o modelo é atualizado periodicamente de acordo com as mudanças físicas. Outra abordagem conhecida é a chamada “*Symbiotic Simulation*” ou “Simulação Simbiótica”. Bergmann, Stelzer e Strassburger (2011) e Onggo *et al.* (2018) relatam que esta abordagem usa dados reais coletados em tempo real ou quase real para atualizar o modelo e

permitir a tomada de decisões por meio de sugestões ou comandos automatizados. Há também a chamada “*Data-driven simulation*” ou “Simulação baseada em dados” que, segundo Meng *et al.* (2013) e Sormaz e Malik (2018), baseia-se na atualização e adaptação periódica do modelo de simulação a partir de dados vindos dos sistemas reais. Por fim, destaca-se também a “Simulação Semi-física”, ou “*Semi-physical Simulation*” como sinônimo dessas abordagens (LENG *et al.*, 2020).

2.2. Indústria 4.0 e o papel dos Gêmeos Digitais

2.2.1. Conceitos e definições

O termo Indústria 4.0, proposto pelo governo alemão em meados de 2011, representa uma revolução dos sistemas produtivos ao fazer alusão ao que seria a quarta revolução industrial (RODIČ, 2017). Diferentemente das três primeiras revoluções industriais, as quais se referem a marcos científicos e inovadores, a quarta revolução industrial representa uma revolução em como as companhias irão reconhecer e lidar com os desafios tecnológicos (RÜTTIMANN e STÖCKLI, 2016). Uriarte, Ng e Moris (2018) relatam que a Indústria 4.0, de maneira simplificada, baseia-se na conexão contínua entre pessoas, sistemas, processos e ambientes, bem como na existência de ambientes físicos e virtuais que se complementam em prol de decisões mais eficientes e eficazes. Neste contexto, todas as informações e dados advindos dos sistemas em geral necessitarão estar totalmente integrados, sendo a conexão entre os ambientes físicos e virtuais a representação de um novo aspecto dos processos produtivos na indústria do futuro (RODIČ, 2017; XU, XU e LI, 2018).

A Indústria 4.0 representa uma meta politicamente estabelecida para os processos produtivos visando integrar e virtualizar os sistemas industriais, englobando neste caso todas as divisões de negócio e da cadeia de agregação de valor (RÜTTIMANN e STÖCKLI, 2016; URIARTE, NG e MORIS, 2018). Neste caso, Rüttimann e Stöckli (2016) afirmam que a Indústria 4.0 é uma tendência inevitável e, dessa forma, as organizações devem se preparar para as mudanças em sua estrutura produtiva. Para os autores, a digitalização dos processos já é perceptível nas mais diversas áreas e segmentos da indústria e, neste caso, soluções tecnológicas quanto à conexão, disponibilidade e ao processamento de dados são peças-chave para o desenvolvimento cada vez mais acelerado dessa nova era. Ainda neste contexto, Lee, Bagheri e Kao (2015) complementam que a maior oferta e disponibilidade de soluções tecnológicas vem impactando a natureza competitiva dos processos produtivos, tornando-se pré-requisitos para a sobrevivência dos mesmos.

Embora do termo Indústria 4.0 tenha sido dedicado inicialmente a atividades ligadas à manufatura (ZHONG *et al.*, 2017), o conceito representa uma revolução em diversos elos da cadeia de suprimentos, englobando atividades relacionadas à logística, qualidade, vendas, segurança, dentre outras (NICOLETTI e PADOVANO, 2019; PRINCES, 2020; SONY, 2018; WINKELHAUS e GROSSE, 2020). Neste caso, espera-se que a era da Indústria 4.0 traga oportunidades de crescimento para todos as frentes da cadeia produtiva, as quais conseguirão sanar grande parte das dificuldades vivenciadas em seus processos e experimentar um novo contexto de operação (RODIČ, 2017). Pereira e Romero (2017) complementam que além da melhoria nos processos associados ao planejamento e produção de bens e serviços, a Indústria 4.0 traz efeitos positivos também na qualidade dos produtos e serviços prestados, no relacionamento entre clientes e organizações, nas relações comerciais, nos resultados financeiros e no ambiente organizacional como um todo.

Ainda quanto aos benefícios trazidos pela Indústria 4.0, Lu (2017) destaca que seus princípios permitem o alcance de um nível mais alto de eficiência operacional e produtividade por parte das organizações. Além disso, Xu, Xu e Li (2018) complementam que tal reformulação dos processos também vai ao encontro do aumento da competitividade imposto aos mesmos. Vale destacar que a Indústria 4.0 deixou de ser um plano futuro e passou a fazer parte do dia-a-dia das organizações. Neste caso, Rodič (2017) ressalta que muitos dos requisitos necessários para a adequação dos processos aos preceitos da Indústria 4.0 já estão disponíveis e cada vez mais acessíveis, sendo a maioria deles de natureza tecnológica. Porém, é importante ressaltar que, ao aderir seus processos aos preceitos da Indústria 4.0, a organização não está somente se tornando mais tecnológica, mas sim uma empresa focada em uma maior agregação de valor e gestão do conhecimento (Lu, 2017). Por outro lado, atrelados aos benefícios trazidos pela Indústria 4.0 estão também os desafios a serem superados pelas organizações (URIARTE; NG; MORIS, 2018), fato que corrobora para a necessidade de pesquisas na área (TORTORELLA; FETTERMANN, 2018).

Segundo Xu, Xu e Li (2018), é possível classificar os desafios associados à Indústria 4.0 em três frentes principais: (i) desafios técnicos; (ii) desafios quanto à padronização das implementações; e (iii) desafios quanto à segurança e privacidade de dados e informações. Os desafios técnicos estão relacionados à infraestrutura atual das empresas, a qual ainda não está preparada para todas as mudanças associadas a este novo contexto industrial. Já quanto aos desafios relacionados à padronização, nota-se certa defasagem na padronização da

implementação de princípios e soluções da Indústria 4.0. Por fim, quanto aos desafios relacionados à segurança e privacidade dos dados e informações, os autores destacam a necessidade de se assegurar a segurança dos dados e informações envolvidas na digitalização dos processos. Além disso, Uhlemann, Lehmann e Steinhilper (2017) alertam para a incerteza acerca dos benefícios econômicos oriundos dessa nova era industrial, frente ao grande investimento necessário para adequar os processos às exigências tecnológicas. Por fim, vale destacar desafios relacionados à provisão de novos conhecimentos e competências requeridas para os funcionários, bem como evolução dos sistemas e práticas de gestão (URIARTE, NG e MORIS, 2018).

2.2.2. Pilares da Indústria 4.0 e o papel dos Gêmeos Digitais

Dentre as principais soluções atreladas à nova era vivenciada pelos processos produtivos, Mueller, Chen e Riedel (2017) destacam: (1) rápida identificação, localização e diagnóstico de parâmetros internos das operações; (2) facilidade na medição do desempenho dos sistemas; (3) acessibilidade aos dados e informações relevantes aos processos; (4) capacidade de comunicar e interagir com objetos e sistemas inteligentes; (5) multifuncionalidade para diferentes aplicações; (6) Observação e interação com os sistemas em ambientes virtualizados; dentre outros. Além disso, tais soluções se baseiam em uma estrutura tecnológica composta por diversas frentes, as quais compõem os pilares da Indústria 4.0 (XU, XU e LI, 2018; ZHONG *et al.*, 2017):

- I. **Internet das coisas:** conhecida em inglês como “*Internet of Things*” (IOT), trata-se de uma tecnologia facilitadora para a próxima geração dos processos produtivos. Segundo Trappey *et al.* (2017), o foco principal da IOT é estabelecer conexões entre os objetos físicos (processos, equipamentos, pessoas, dentre outros) com uma rede de internet onipresente, possibilitando a coleta e troca de dados e informações entre tais objetos de maneira contínua. Manavalan e Jayakrishna (2019) complementam que a IOT, por meio do compartilhamento de dados em tempo real ou quase real, vem ganhando espaço e já está presente em diversos setores da cadeia produtiva;
- II. **Tecnologia de nuvem:** conhecida em inglês como “*Cloud*” ou “*Cloud Computing*”, tal tecnologia permite o compartilhamento, alocação e processamento de dados e informações a partir de redes e servidores, tudo de maneira flexível e facilitando as operações (XU, XU e LI, 2018). Os autores ressaltam que as decisões são cada vez mais baseadas em dados e ferramentas de análise que requerem múltiplos recursos

computacionais. Neste caso, ao se considerar o uso da tecnologia de nuvem como suporte, têm-se um processo decisório mais eficiente. A tecnologia de nuvem também contribui para o contexto da Indústria 4.0 ao permitir o compartilhamento e processamento de dados entre as fronteiras da empresa, melhorando o desempenho do sistema produtivo, tornando-o mais ágil e flexível, bem como contribuindo para a redução de custos associados ao armazenamento e processamento de dados e informações (LIU e XU, 2017).

- III. *Big data*:** para Qi e Tao (2018), o termo em inglês “*Big data*” refere-se a técnicas de obtenção, processamento e análise de dados em grande quantidade e variedade, estruturados ou não, de forma a se obter informações a partir dos mesmos. De acordo com os autores, os processos produtivos estão inseridos em um ambiente marcado por um grande volume de dados estruturados, não estruturados e semiestruturados, os quais são gerados a partir do ciclo de vida dos produtos, serviços e processos. Neste contexto, a partir do uso de técnicas, modelos e ferramentas inseridas no contexto *Big Data*, os tomadores de decisão podem encontrar ineficiências dos processos, identificar as causas e impactos dos problemas, analisar comportamentos, prever soluções, dentre outras facilidades;
- IV. Tecnologia de informação (TI):** segundo Zhong *et al.* (2017), as tecnologias de informação, ou também conhecidas como “*Information Technology*” em inglês, referem-se a toda estrutura capaz de permitir o processamento e transferência de sinais, dados e informações entre diversos sistemas. Neste caso, apesar de não se tratar de uma nova tecnologia ou solução, a estrutura de TI vem evoluindo ao longo do tempo se potencializa no âmbito da Indústria 4.0, sendo combinada com as soluções de IOT e *cloud*;
- V. Sistemas Ciber-físicos:** também chamados de “*Cyber-physical Systems*” (CPS) em inglês, referem-se à junção entre os sistemas físicos e virtuais, os quais são conectados e sincronizados por meio da estrutura de TI, compondo um ambiente integrado, com autonomia parcial ou completa, e capaz de cooperar com a interface humana para a tomada de decisões (MUELLER, CHEN e RIEDEL, 2017). Além disso, os autores complementam que outra característica fundamental do CPS é o grau de inteligência utilizado para apoio à decisão. Logo, não se trata somente da virtualização de processos, mas também da potencialização dos mesmos por meio de técnicas e sistemas auxiliares que compõem o ambiente ciber-físico. Neste contexto, Shin, Cho e Oh (2018) relatam que o CPS desempenha um papel crucial na concepção da Indústria 4.0, uma vez que

permite a digitalização dos sistemas físicos que compõem os processos produtivos (sensores, atuadores, dispositivos móveis e máquinas, dentre outros). Neste caso, há o espelhamento do comportamento desses sistemas durante sua operação com foco no gerenciamento síncrono.

Alguns autores citam outras soluções como sendo elementos-chave da Indústria 4.0: sistemas adaptáveis, flexíveis e com certa automação; segurança na comunicação; Inteligência Artificial (IA); sistemas robóticos colaborativos; realidade virtual; dentre outros (LONGO e NICOLETTI; PADOVANO, 2017; LU, 2017; MOEUF *et al.*, 2018; XU, XU e LI, 2018; YIN, STECKE e LI, 2018). Entretanto, nota-se que a maioria dessas soluções estão inseridas de certa forma nos cinco pilares descritos. Além disso, vale ressaltar que o CPS se destaca como o pilar-chave para a difusão dos princípios da Indústria 4.0, já que baseia-se nos demais para sua concepção (MUELLER, CHEN e RIEDEL, 2017). Outro ponto importante diz respeito à estrutura que compõe o CPS. Shin, Cho e Oh (2018) sugerem uma visão baseada em cinco níveis: (i) nível de conexão (nível mais básico onde ocorre a obtenção dos dados físicos por meio de sensores, equipamentos, sistemas de gestão, dentre outros); (ii) nível de conversão (responsável pela conversão dos dados brutos em informações); (iii) nível cibernético (no qual as informações são transferidas para o nível cibernético, permitindo a estruturação do modelo virtual); (iv) nível cognitivo (permite uma visualização remota do sistema físico para apoio à tomada de decisões); e (v) nível de configuração (onde o ocorre o *feedback* do ambiente virtual para os sistemas físicos).

Além disso, ainda com relação à representação virtual dos sistemas físicos, destaca-se o chamado “Gêmeo Digital” ou em inglês, “*Digital Twin*”. Boschert e Rosen (2016) simplificam o conceito de GD como sendo uma representação virtual que espelha o comportamento de um componente, produto ou sistema e permite sua avaliação durante todo o ciclo de vida. Neste caso, conclui-se que os GDs representam a virtualização dos sistemas físicos, à qual se refere o CPS, atuando como um pré-requisito fundamental para a concepção desses sistemas (KUNATH e WINKLER, 2018; UHLEMANN, LEHMANN e STEINHILPER, 2017). O GD, em sua primeira definição, foi descrito como uma ferramenta de simulação multi-física, multi-escala e probabilística de um sistema, o qual coleta dados e espelha os comportamentos físicos por meio de um modelo virtual. Além disso, trata-se de um modelo capaz de avaliar e recomendar mudanças para os sistemas físicos, a fim de otimizá-los (SHAFTO *et al.*, 2010). Entretanto, desde que o GD foi criado, nota-se diferentes definições e abordagens no decorrer dos anos.

Segundo Kritzinger *et al.* (2018), apesar de ser reconhecido inicialmente como um espelhamento virtual de um produto ou processo, o GD é mais que puramente uma cópia virtual, já que contempla técnicas e ferramentas capazes de avaliar o sistema real e orientar decisões baseadas no seu comportamento. Neste sentido, Tao *et al.* (2018) destacam como sendo as principais características do GD, além da capacidade de espelhar o sistema físico de maneira altamente sincronizada e fiel, a capacidade de se conectar com diversas fontes de dados e reconhecer as características e comportamento dos sistemas ao longo do tempo de forma a se obter decisões cada vez mais eficientes e assertivas. Em contraste com decisões baseadas na experiência e expertise dos tomadores de decisões, Kunath e Winkler (2018) salientam que a era dos GDs baseia-se em decisões cada vez mais eficientes e independentes do decisor. Por fim, é importante ressaltar que, considerando os preceitos da Indústria 4.0, entende-se que os sistemas produtivos devem ser capazes de monitorar e coletar informações de seus processos de forma a abastecer os GDs que, por sua vez, cooperam com os tomadores de decisões (ZHONG *et al.*, 2017).

Rodič (2017) relata que o GD é composto por dois elementos principais, chamados de “*Digital Shadow*” e “*Digital Master*”, ou em português “Sombra Digital” e “Controlador Digital”, respectivamente. Para o autor, a sombra digital refere-se aos dados advindos dos sistemas físicos enquanto que o controlador digital se baseia na tradução destes dados em modelos e representações que permitem o espelhamento do ambiente físico. De maneira mais detalhada, Tao e Zhang (2017) relatam que, embora haja diferentes abordagens do GD, as quais variam de acordo com o nível de integração entre os ambientes físico e virtual, todos podem ser compreendidos a partir de quatro componentes principais: (i) Sistema Físico (SF), composto por processos, máquinas, equipamentos, dentre outros; (ii) Sistema Virtual (SV), que consiste em modelos que representam o comportamento do sistema físico ao longo do tempo; (iii) Sistema de Serviço (SS), o qual inclui a estrutura capaz de permitir a comunicação entre os ambientes físico e virtual; e (iv) Sistema de Dados (SD), o qual representa o conjunto de dados e informações que são transmitidos entre os sistemas.

Além disso, muito se discute a respeito dos requisitos do GD no que diz respeito ao nível de sincronismo entre os ambientes físico e virtual e o grau de autonomia na tomada de decisões. Com relação ao sincronismo, o modelo virtual deve ser atualizado com um atraso mínimo de forma a não se prejudicar a tomada de decisão e, dessa forma, entende-se que abordagens em tempo real e quase real são válidas (KUNATH e WINKLER, 2018; UHLEMANN, LEHMANN e STEINHILPER, 2017). Já com relação ao grau de autonomia do modelo virtual, a literatura apresenta alguns exemplos de GDs autônomos, onde os modelos são capazes de controlar os

sistemas físicos sem intervenção humana (BEREGI, SZALLER e KÁDÁR, 2018; DONHAUSER *et al.*, 2018). No entanto, é importante ressaltar que as respostas do GD podem ser no formato de recomendações e diretrizes para a tomada de decisão. Neste caso, o modelo virtual auxilia na obtenção das melhores decisões e fornece orientações ao tomador de decisões (ALAM e SADDIK, 2017; LU *et al.*, 2019; VIJAYAKUMAR *et al.*, 2019).

Desde que foi criado, o conceito do GD vem sendo aperfeiçoado a partir de aplicações das mais diversas áreas. Wright e Davidson (2020) relatam aplicações em processos industriais, sistemas de saúde, área de serviços, contemplando os mais diversos escopos, dentre outras. Além disso, os autores destacam GDs tanto de produtos, como motores, máquinas e equipamentos, como de processos, contemplando o fluxo dos sistemas produtivos. Pode-se ainda entender as características de decisão baseadas em GDs de acordo com a fase dos processos na qual os modelos são utilizados (antes e/ou durante a operação dos processos). Por um lado, o uso de GD para guiar as decisões antes operação de processos está ligada à fase de planejamento e design, onde pode-se testar cenários e verificar a efeito das decisões antes de implementá-las, validando o sistema físico (LIU *et al.*, 2019; LIU *et al.*, 2021). Sobre por outro lado, ao considerar o uso de DT durante a operação dos processos, pretende-se realizar um estudo paralelo que envolve a avaliação e controle dos sistemas físicos, bem como o suporte à tomada de decisões durante todo o ciclo de vida dos processos (LENG *et al.*, 2019; ALAM e SADDIK, 2017).

Ainda com relação aos GDs de processos, Tao e Zhang (2017) destacam que estes representam um importante marco para os sistemas produtivos, trazendo melhores decisões por meio do monitoramento e controle dos sistemas físicos. Por fim, segundo Alam e Saddik (2017), o GD pode ser usado para apoiar decisões a partir de três abordagens principais: (i) diagnóstico, visando avaliar decisões com base em análises acerca do passado; (ii) monitoramento, a fim de monitorar e controlar os processos; e (iii) prognóstico, visando antecipar e prever comportamentos. A Figura 2.2 ilustra a estrutura genérica do GD de forma a sumarizar suas principais características e componentes.

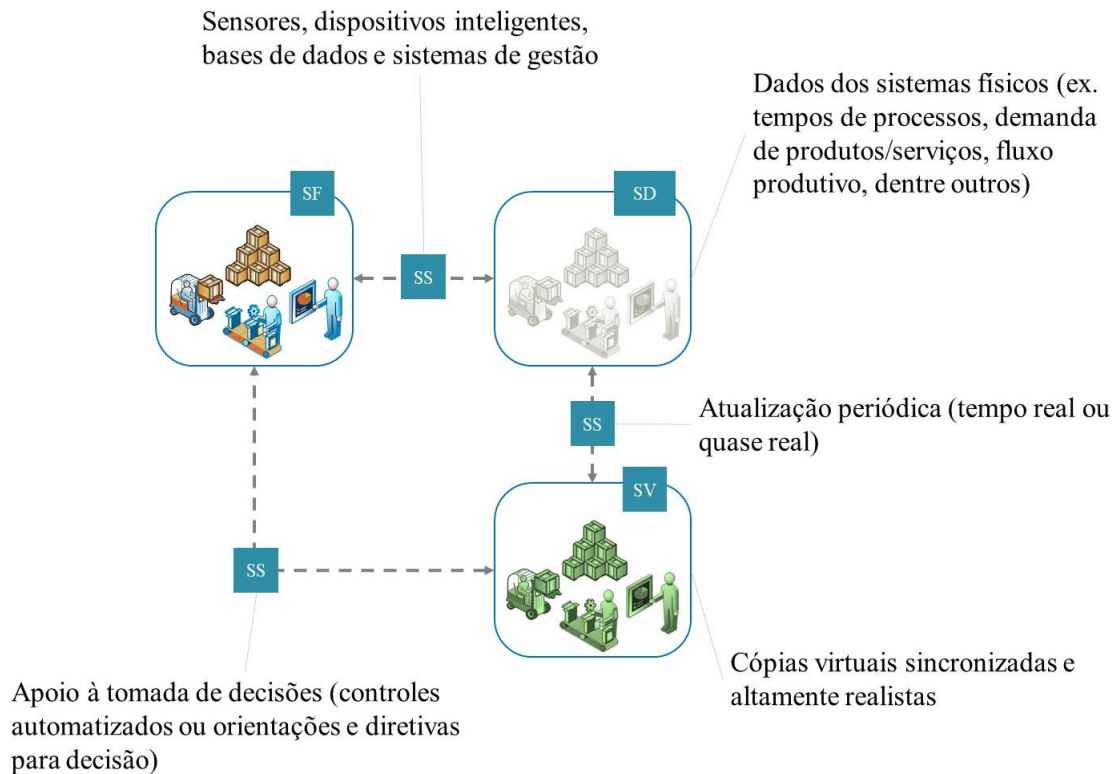


Figura 2.2 – Estrutura geral do Gêmeo Digital para apoio à decisão

Fonte: elaborado pelo próprio autor

2.3. Uso da simulação como Gêmeo Digital de processos

2.3.1. Conceitos e definições

Conforme destacado anteriormente, a simulação, considerando abordagens tradicionais, tende a ser incompatível com o caráter dinâmico dos processos produtivos e com as necessidades relacionadas à tomada de decisão. Além disso, com base na evolução da simulação no decorrer dos anos, Mourtzis (2020) relata que os sistemas produtivos vivenciam uma nova era da simulação computacional. Neste contexto, tendo em vista o ambiente impulsionado pela Indústria 4.0 e seus principais pilares e soluções, a simulação se mostra capaz de assumir o papel dos GDs (SANTOS *et al.*, 2022a; TIACCI, 2020; ZHUANG, LIU e XIONG, 2018). Tal afirmação se justifica já que o uso de modelos de simulação como GDs acabou se tornando uma alternativa às soluções comerciais, as quais geralmente acompanham equipamentos e sistemas prontos oferecidos pelas grandes marcas industriais (TERKAJ *et al.*, 2019). Neste caso, o uso da simulação para tal fim representa uma abordagem que costuma envolver menores investimentos, já que não está associada à aquisição de equipamentos, e também maior flexibilidade, sendo compatível com processos produtivos automatizados ou não e com diferentes escopos (SANTOS *et al.*, 2022a; TERKAJ *et al.*, 2019).

De acordo com Wright e Davidson (2020), o que difere um modelo tradicional para um GD é a capacidade de estender seu uso para escalas de tempo sobre a qual o sistema físico mudará constantemente. Logo, um modelo de simulação capaz de se conectar aos sistemas físicos e se adaptar de acordo com suas mudanças pode ser classificado como GD (ASHRAFIAN *et al.*, 2019; SANTOS *et al.*, 2022a). O quadro 2.1 ilustra as principais diferenças entre a abordagem tradicional da simulação e o GD. Além disso, embora altamente promissor, a tomada de decisões baseada em GDs ainda não é uma área totalmente explorada (MOURTZIS, 2020). Isso porque ainda há questões em aberto sobre as características de um GD e, ao se considerar a simulação como tal, essas questões devem ser igualmente abordadas. Dentre os principais pontos, destaca-se o tipo de simulação adotado, o intervalo de tempo entre as atualizações do modelo, etapas associadas à construção do modelo, segurança relacionada a tomada de decisão, o nível de detalhe e integração dos modelos com os sistemas físicos, a necessidade de avaliar periodicamente a validade do modelo, dentre outros (SANTOS *et al.*, 2022a; TAO e ZHANG, 2017; WRIGHT e DAVIDSON, 2020).

Quadro 2.1 – Simulação em aplicações tradicionais vs. Gêmeo Digital

Característica	Simulação tradicional	Simulação como Gêmeo Digital
Coleta de dados	Manual	Automática
Atualização do modelo	Manual	Automática e periódica
Percepção de tempo	Análises pontuais	Análises periódicas (tempo real ou quase real)
Escopo de atuação	Limitado	Amplo e flexível
Principais objetivos	Análises específicas	Apoio constante à tomada de decisões
Complexidade	Necessidade de especialistas	Modelos amigáveis
Integração com os sistemas físicos	Não aplicável	Parcial ou total
Autonomia	Não autônoma	Autônoma ou semiautônoma (diretivas para decisão)

Fonte: adaptado de Santos *et al.* (2022a)

Nota-se que a literatura apresenta diversas aplicações envolvendo o uso da simulação como GD, com destaque para publicações nos últimos anos (BEREGI, SZALLER e KÁDÁR, 2018; GRUBE, MALIK e BILBERG, 2019; KARAKRA *et al.*, 2018; MURPHY *et al.*, 2020; PRAJAPAT *et al.*, 2019; STERINGER *et al.*, 2019; TERKAJ *et al.*, 2019; VIJAYAKUMAR *et al.*, 2019). Por outro lado, embora o GD seja um conceito relativamente recente, conectar modelos de simulação aos sistemas físicos e atualizá-los ao longo do tempo é uma prática que vem sendo adotada há décadas. Em 1993, Katz e Manivannan (1993) já utilizavam a simulação conectada com equipamentos de uma linha de manufatura visando monitorá-los. Logo, apesar de não utilizar o termo “Gêmeo Digital”, nota-se que várias nomenclaturas têm sido usadas ao longo dos anos para se referir a essa abordagem: “Simulação simbiótica”; “Simulação *online*”; “Simulação orientada a dados”; “Simulação em tempo real”; e “Simulação quase em tempo real” (ONGGO *et al.*, 2018; SAEZ *et al.*, 2018; SCHOLL *et al.*, 2012; SORMAZ e MALIK, 2018; VAHDATIKHAKI e HAMMAD, 2014). Outro ponto importante a ser abordado diz respeito à evolução desta abordagem nos últimos anos. Espera-se que algumas dificuldades relatadas no passado já tenham sido superadas devido aos avanços tecnológicos e, por outro lado, espera-se também novos desafios e oportunidades a serem exploradas.

2.3.2. Validade dos modelos de simulação utilizados como Gêmeos Digitais

Com base nas observações destacadas até aqui, reforça-se a importância de trabalhos focados na garantia da validade dos modelos de simulação utilizados como GD. Neste contexto, Tao e Zhang (2017) revelam que, para tal, são necessárias rotinas periódicas de Verificação, Validação e Acreditação (VV&A) do modelo. Neste caso, enquanto as etapas de verificação e validação buscam garantir o correto funcionamento do modelo computacional e sua correspondência satisfatória com os sistemas físicos, respectivamente, a fase de acreditação está relacionada ao tomador de decisão e avalia a confiabilidade do modelo para orientar decisões (SARGENT, 2011). Apesar da importância de se garantir tais rotinas de VV&A durante a operação do GD, Zhuang, Liu e Xiong (2018) e Onggo *et al.* (2018) revelam que ainda faltam métodos e técnicas para esse fim. Além disso, é importante ressaltar que, considerando o uso da simulação como GD, há duas fases distintas e complementares que devem ser consideradas sob o ponto de vista da validade do modelo: (i) construção do modelo e (ii) operação do modelo (MONTEVECHI *et al.*, 2020).

Considerando a validade do modelo de simulação durante sua fase de construção do GD, nota-se que procedimentos análogos ao de abordagens tradicionais de simulação podem ser adotados

(MONTEVECHI *et al.*, 2020; SANTOS *et al.*, 2022a). Tal afirmação é válida uma vez que grande parte das etapas consideradas na construção do modelo são semelhantes, independente da abordagem adotada. Neste contexto, a literatura apresenta um vasto acervo de métodos e técnicas para garantir a validade do modelo em sua construção, chamados, de maneira simplificada, de métodos de validação. De uma maneira geral, pode-se classificar os métodos de validação em objetivos e subjetivos (SARGENT, GOLDSMAN e YAACOUB, 2016). Neste caso, as técnicas objetivas são baseadas em procedimentos matemáticos e estatísticos visando garantir a validade do modelo, enquanto métodos subjetivos são baseados na experiência e percepção dos modeladores. Leal *et al.* (2011) revelam que somente a adoção de métodos subjetivos dificilmente é capaz de atestar a validade dos modelos e, neste contexto, métodos objetivos são desejáveis sempre que possível. Logo, técnicas subjetivas são úteis como análises complementares ou quando não é possível a adoção de técnicas objetivas (LEAL *et al.*, 2011; WANG, 2013).

De maneira mais detalhada, Sargent (2013) descreve os principais métodos de validação do modelo de simulação durante sua fase de construção. Dentre as principais técnicas subjetivas, pode-se citar: (i) Animação (o comportamento do modelo é avaliado graficamente); (ii) Comparação (onde o comportamento do modelo é comparado com outros modelos ou resultados históricos); (iii) Teste de condições extremas (onde o modelo é avaliado a partir da indução de comportamentos extremos); (iv) Validação face-a-face (no qual especialistas são consultados acerca do comportamento do modelo); (v) Análise de sensibilidade (onde alguns parâmetros do modelo são alterados de forma a se avaliar o comportamento desejado); (vi) Análise de gráficos (onde gráficos, como *boxplots* e histogramas, são utilizados visando comparar resultados reais com simulados); dentre outros. Já ao se considerar métodos objetivos, destacam-se basicamente testes e análises estatísticas, as quais podem se basear em: (vii) Testes de hipóteses (testes estatísticos que comparam variáveis do modelo com suas correspondentes reais); e (viii) Análise de intervalo de confiança (onde a faixa de precisão do modelo é avaliada em função de um intervalo desejável).

Por outro lado, ao se considerar a avaliação da validade dos modelos de simulação durante sua operação como GD, não há métodos propostos na literatura. Os únicos autores que propuseram tal prática, CHO *et al.* (2019), adotaram rotinas de testes de hipótese na tentativa de comparar sistemática e periodicamente o comportamento do modelo com o ambiente físico. Mais detalhes acerca dessa abordagem, bem como sobre o estado da arte sobre o tema são explorados em mais detalhes na RSL apresentada no capítulo 4 deste trabalho. Neste caso, entende-se que tal abordagem é válida, porém há algumas questões que podem dificultar ou mesmo invalidá-la. O

primeiro ponto a se destacar está relacionado à complexidade do GD. No caso de modelos complexos, conectados à diversos sistemas físicos, têm-se um número considerável de variáveis a se avaliar de forma a garantir a validade do modelo e, neste caso, seriam necessários diversos testes de hipóteses simultâneos. Outro ponto a se considerar diz respeito ao comportamento dessas variáveis. No caso do teste de hipótese, considera-se estatísticas descritivas (média e desvio padrão, por exemplo) para descrever cada variável avaliada, porém há o risco de se ter diferentes comportamentos a partir das mesmas estatísticas descritivas (CHEN *et al.*, 2019). Por fim, a validade do modelo pode se alterar gradativamente no decorrer de seu uso e, ao considerar apenas testes de hipóteses periódicos, desconsidera-se tal ocorrência no decorrer do tempo.

Portanto, conclui-se a importância de se desenvolver trabalhos focados em técnicas e métodos de avaliação da validade dos modelos de simulação durante sua operação do GD. Neste caso, nota-se que as abordagens utilizadas na avaliação da validade durante a fase de construção dos modelos não contemplam certas características necessárias para sua utilização durante a operação do GD, conforme destacado anteriormente. Portanto, este trabalho propõe uma abordagem baseada na utilização do K-NN, uma técnica de Aprendizado de Máquina, visando avaliar a validade dos modelos de simulação como GD. Tal técnica possui a capacidade de avaliação simultânea de diversas variáveis, inclusive de natureza diferente quanto ao seu tipo e comportamento, tornando possível a avaliação da validade do GD. Além disso, uma vez que o modelo possui caráter dinâmico ao se adaptar às mudanças físicas, é necessária a avaliação periódica de sua validade e, neste contexto, propõe-se a adoção combinada do K-NN com a carta de controle. Dessa forma, espera-se obter uma ferramenta de monitoramento capaz de assegurar a validade do modelo no decorrer do tempo. As seções seguintes são focadas na apresentação de ambas as técnicas adotadas neste trabalho.

2.4. Aprendizado de Máquina e o *K-Nearest Neighbors* (K-NN)

2.4.1. Conceitos e Definições

As técnicas de Aprendizado de Máquina (AM), do inglês *Machine Learning*, compõem uma das áreas de IA que despontam no cenário da Indústria 4.0 e dizem respeito ao desenvolvimento e aplicação de algoritmos computacionais capazes de aprender certos padrões e relações complexas, baseando-se em informações e dados, sendo amplamente utilizadas para apoio à decisão (NI, XIAO e LIM, 2020). Avci *et al.* (2021) destacam que a base das técnicas de AM é fornecer ao computador uma habilidade semelhante à humana em termos de aprendizado,

tornando-o capaz de aprender com comportamentos passados e apoiar às decisões a partir de sua experiência adquirida.

Para Sharma *et al.* (2020), uma estrutura geral de AM é baseada primeiramente em um conjunto de dados de treinamento (rotulados ou não) que abastecem o algoritmo de AM, o qual aprende os comportamentos e padrões e, a partir dessa experiência, é capaz de inferir acerca de novos dados (dados de teste) inseridos no sistema. A Figura 2.3 ilustra tal estrutura. Além disso, as técnicas de AM podem ser classificadas em: (i) Supervisionadas (onde os dados de treinamento do algoritmo são rotulados e, dessa forma, ensina-se o algoritmo os comportamentos e padrões a serem avaliados); e (ii) Não supervisionadas (nas quais os dados de treinamento não são rotulados e o algoritmo aprende a partir das semelhanças e características desses dados) (PRIYADHARSHINI e JOSEPHIN, 2020).

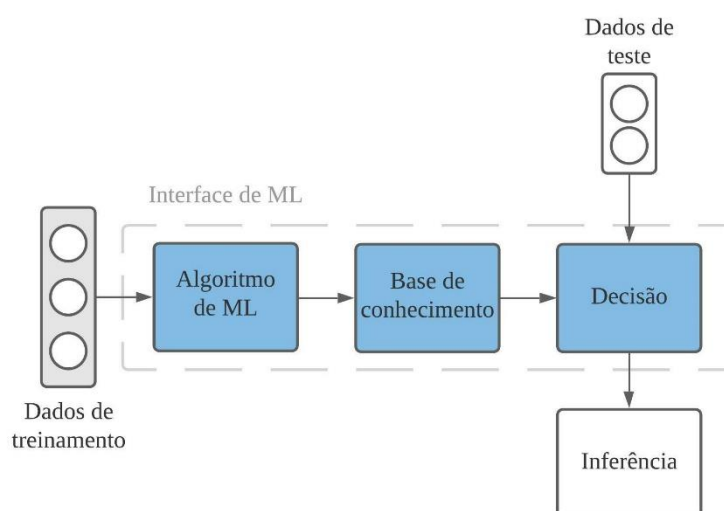


Figura 2.3 – Estrutura geral das técnicas de AM

Fonte: Adaptado de Sharma *et al.* (2020)

Com relação aos objetivos atrelados aos estudos de AM, Avci *et al.* (2021) destacam como sendo quatro principais: (i) Classificação (onde os dados de teste recebem rótulos referente a categorias previamente definidas); (ii) Regressão (cujo objetivo é associar as saídas com as entradas de forma a se obter uma relação entre ambas); (iii) Previsão (trata-se de um tipo especial de regressão focado na previsão de valores futuros); e (iv) Agrupamento (onde os dados são agrupados de acordo com suas características). Neste caso, os autores ressaltam que enquanto a Classificação, Regressão e Previsão são métodos supervisionados, o Agrupamento trata-se de um método não supervisionado.

Já com relação aos principais algoritmos de AM adotados na literatura, pode-se citar: Redes Bayesianas, Árvore de Decisão, Análises de Regressão, *Support Vector Machine*, Redes Neurais Artificiais, algoritmos de *Ensemble*, Aprendizado Profundo, Algoritmo Genético e Aprendizado Baseado em Instância (SHARMA *et al.*, 2020). Além disso, os autores dão destaque para o *K-Nearest Neighbors* (K-NN) como uma valiosa técnica que integra os algoritmos de Aprendizado Baseado em Instância. Ni, Xiao e Lim (2020) revelam que o K-NN costuma ser amplamente utilizado para objetivos relacionados à Regressão, Classificação e Agrupamento devido a sua baixa complexidade atrelada a bons resultados.

2.4.2. K-NN como classificador

O K-NN é uma técnica de AM amplamente difundida na literatura e sua utilização como classificador possui grande potencial devido às suas características (BOONTASRI e TEMDEE, 2020; NI, XIAO e LIM, 2020). Segundo Lee *et al.* (2020), o K-NN é um método supervisionado de AM e figura como um dos classificadores mais simples e rápidos, fato que colabora para sua ampla utilização. Além disso, sua disseminação nos últimos anos se deu também devido à sua robustez e adequação para problemas com várias classes de dados e com conjunto de dados onde há pouco ou nenhum conhecimento prévio do seu comportamento (KUMAR *et al.*, 2020; ZHANG *et al.*, 2021).

Embora haja dificuldades com sua aplicação no caso de dados faltantes e não regulares, Ni *et al.* (2019) revelam que outra vantagem associada ao K-NN é o fato de se tratar de uma técnica imune a ruídos, ou seja, pontos considerados fora do contexto (*outliers*) que podem prejudicar as análises. O K-NN se baseia na atribuição de rótulos aos dados de teste com base nos rótulos dos dados de treinamento mais próximos, ou também chamados de vizinhos. Neste caso, o algoritmo irá se basear em um número ‘K’ de vizinhos mais próximos (KUMAR *et al.*, 2020). Para a definição dos vizinhos, o K-NN baseia-se na chamada “distância euclidiana”, a qual é obtida a partir da equação 1 (GONG, SU e TSENG, 2020). Vale ressaltar que, caso os dados sejam de natureza diferente, é necessária a padronização dos mesmos antes de considerar tal distância (GHOSH, 2006).

$$D = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

Onde D representa a distância entre dois vetores de n dimensões, $\mathbf{X} = (x_1, x_2, \dots, x_n)$ e $\mathbf{Y} = (y_1, y_2, \dots, y_n)$.

Apesar da grande aplicabilidade, Ghosh (2006) destaca que um dos grandes desafios relacionados ao uso do K-NN refere-se à escolha do parâmetro 'K', uma vez que o desempenho do classificador depende diretamente desse valor. Neste caso, o valor 'K' irá variar de acordo com as características da aplicação, e escolher o valor ótimo é uma tarefa difícil, pois depende do conjunto de dados adotado (HASSANAT, ABBADI e ALHASANAT, 2014; SOLHMIRZAEI *et al.*, 2020). Zhang *et al.* (2021) afirmam que um valor de 'K' maior que o necessário pode melhorar a capacidade do modelo em compreender os padrões, porém aumenta a complexidade do algoritmo. Por outro lado, um valor de 'K' menor que o desejável pode prejudicar a capacidade do modelo em compreender os padrões, ocorrendo o chamado *overfitting*, que se refere a um estado onde o modelo é enviesado pelos dados de treinamento e não consegue obter conhecimento suficiente para classificar os dados de teste. A Figura 2.4 ilustra o procedimento de classificação de K-NN e a importância da escolha do valor de 'K'. Nesse caso, para $K = 3$, notamos que os vizinhos mais próximos possuem em sua maioria "Rótulo B", enquanto que para $K = 7$, o "Rótulo A" é predominante.

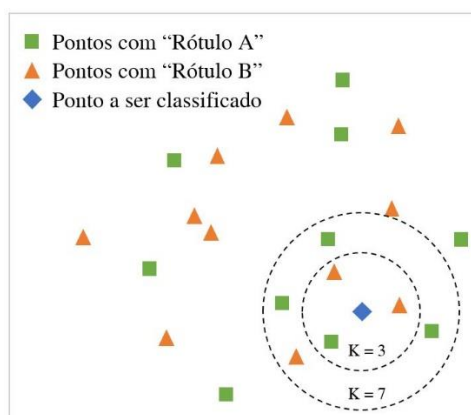


Figura 2.4 – Influência do valor de 'K' na classificação pelo K-NN

Fonte: Adaptado de Zhang *et al.* (2021)

Outro fator que impacta o desempenho do K-NN é o procedimento adotado para treinamento do classificador, uma vez que se trata de um método supervisionado (WANG *et al.*, 2020). Neste contexto, é necessário dividir os dados entre treinamento e teste e uma prática comum na literatura é a divisão 70%/30%, respectivamente (HASSANAT, ABBADI e ALHASANAT, 2014; WANG *et al.*, 2020). Entretanto, há a possibilidade de haver um viés durante o processo

de treinamento e, neste caso, uma solução amplamente adotada é a chamada *Cross-validation* ou Validação Cruzada (SOLHMIRZAEI *et al.*, 2020). Neste caso, os dados são divididos em ‘Z’ subconjuntos e, enquanto um subconjunto é utilizado como dados de teste, os demais (Z-1) são utilizados como dados de treinamento. Além disso, tal processo é repetido ‘Z’ vezes de forma a se obter um modelo mais eficiente (BOONTASRI e TEMDEE, 2020). Dessa forma, a partir de uma boa escolha do valor de ‘K’ e do treinamento correto do algoritmo, espera-se um bom desempenho do classificador.

O desempenho do K-NN é frequentemente medido a partir de sua acurácia (BOONTASRI e TEMDEE, 2020). Neste caso, tal acurácia é obtida a partir da taxa de acertos do classificador, e, dessa forma, uma precisão de 100% indica que o K-NN é capaz de classificar perfeitamente os dados a partir de seus vizinhos próximos. Por outro lado, caso o K-NN não consiga identificar de maneira clara os rótulos vizinhos, o algoritmo atribui uma classificação aleatória ao ponto e, neste caso, a acurácia tende a 50% (KUMAR *et al.*, 2020). Neste contexto, o intuito deste trabalho é associar a acurácia do K-NN com a validade do GD. A partir da coleta de dados dos sistemas físicos e virtuais, objetiva-se testar a capacidade do classificador em diferenciar ambos e, caso não consiga, têm-se um modelo virtual semelhante ao sistema real, fato que atesta sua validade. Dessa forma, objetiva-se a acurácia do K-NN por volta de 50%. Além disso, outro ponto importante é que o GD, conforme destacado anteriormente, possui um caráter dinâmico e, dessa forma, é necessário avaliar sua validade no decorrer do tempo. Logo, para tal, optou-se pela adoção da carta de controle como mecanismo de monitoramento.

2.5. Cartas de controle como mecanismo de monitoramento

2.5.1. Conceitos e Definições

De acordo com Montgomery (2009), as cartas de controle são uma das principais técnicas de controle estatístico de processo, as quais são úteis para se monitorar variáveis de saída em sistemas sujeitos a fontes indesejáveis de variabilidade. Propostos inicialmente por Walter Shewhart, os gráficos de controle são usados para sinalizar a presença dessas fontes de variabilidade. Basicamente, amostras são regularmente retiradas do processo e o valor de uma estatística de monitoramento é calculado (SHEWHART, 1926). Em seguida, essa estatística é plotada em um gráfico e, caso o ponto seja plotado externamente aos limites de controle do gráfico, o processo é considerado fora de controle e ações corretivas devem ser tomadas (APSEMIDIS, PSARAKIS e MOGUERZA, 2020; ZWETSLOOT e WOODALL, 2021). Neste caso, uma carta de controle é considerada mais eficiente na medida em que é capaz de notificar

mais rapidamente as variações incomuns no processo (ABBAS *et al.*, 2019). Desde sua proposição, as cartas de controle vêm sendo amplamente utilizadas nas mais diversas áreas, como manufatura, saúde e serviços e, considerando as particularidades dessas áreas, diferentes tipos de cartas têm sido desenvolvidos (ABBAS *et al.*, 2019; CHUKHROVA e JOHANNSEN, 2019).

É importante ressaltar que o estudo envolvendo o uso das cartas de controle pode ser compreendido a partir de duas fases principais, chamadas de “Fase I” e “Fase II” (CHUKHROVA e JOHANNSEN, 2019; ZWETSLOOT e WOODALL, 2021). A fase I compreende a etapa onde os parâmetros da carta são definidos, isto é, o Limite Superior de Controle (LSC), Limite Inferior de Controle (LIC) e a Linha Central (LC) (SABAHNO, CASTAGLIOLA e AMIRI, 2020). Neste caso, deve-se garantir que o processo esteja sob controle e em condições normais de operação. A Figura 2.5 ilustra de maneira genérica a estrutura da carta de controle. Já a fase II do estudo refere-se à etapa de monitoramento do sistema com base na carta construída previamente, ou seja, refere-se diretamente ao objetivo do estudo, que é a avaliação do sistema ao longo do tempo (ZWETSLOOT e WOODALL, 2021). Além disso, diversas abordagens exploram o monitoramento das características de qualidade que são expressas em uma escala quantitativa e, neste caso, são classificadas como “Cartas de Variáveis”. Por outro lado, algumas características de qualidade são expressas como variáveis qualitativas e, neste contexto, tem-se as chamadas “Cartas de Atributos” (MONTGOMERY, 2009).

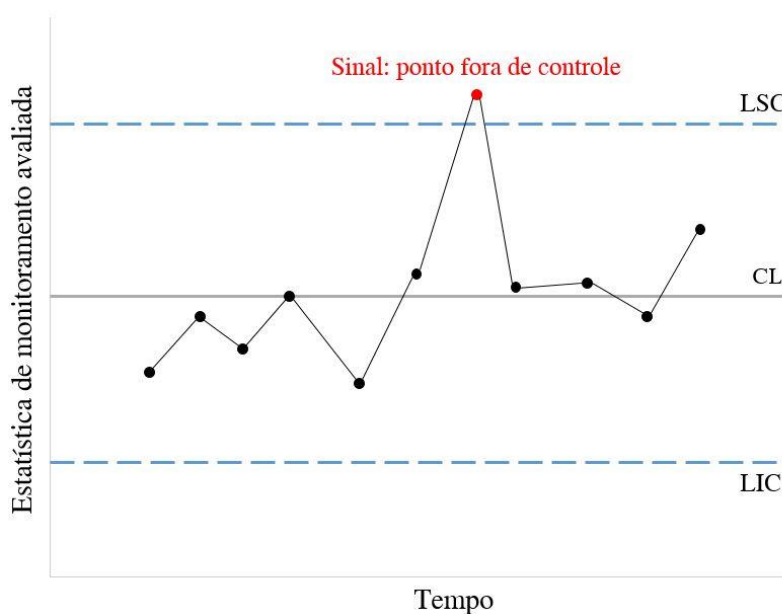


Figura 2.5 – Estrutura geral da carta de controle

Fonte: Adaptado de Montgomery (2009)

Considerando cartas de variáveis mais populares, pode-se destacar: (i) Carta \bar{X} (focada no monitoramento da média amostral); (ii) Carta R (que avalia a amplitude amostral); (iii) Carta S (focada na avaliação da variância amostral); (iv) Carta X (com foco em observações individuais); (v) Carta MR (focada na amplitude móvel); dentre outras. Já com relação às cartas de atributos, destacam-se: (vi) Carta np (focada no número de “itens” com uma certa característica); Carta p (com foco na proporção de “itens” com uma certa característica); Carta C (que avalia o total de uma certa característica presente na amostra); dentre outras (COSTA, EPPRECHT e CARPINETTI, 2018; MONTGOMERY, 2009). Costa, Epprecht e Carpinetti (2018) destacam a importância das cartas de variáveis nos últimos anos, fornecendo amplo apoio no monitoramento de processos. Entretanto, as cartas de atributos têm sido amplamente exploradas por pesquisadores devido à sua simplicidade e praticidade, podendo ser inclusive uma alternativa economicamente viável para monitoramento de processos se comparada com as cartas de variáveis (LEONI e COSTA, 2018; ZHOU, CHENG e ZHENG, 2019).

Considerando a acurácia do K-NN como parâmetro do processo a ser monitorado, optou-se por adotar a carta de controle de atributo p para este trabalho. Tal escolha se justifica uma vez que tal acurácia pode ser obtida considerando a proporção de erros e acertos do classificador ao tentar separar os dados físicos dos virtuais. Logo, espera-se integrar a carta p ao classificador K-NN de forma a compor uma ferramenta de controle robusta, capaz de avaliar periodicamente a acurácia do K-NN e, conseqüentemente a validade do GD durante sua operação.

2.5.2. Carta de controle p

Para Chukhrova e Johannssen (2019), a carta de controle p vem sendo amplamente utilizada nas mais diversas áreas de aplicação. Montgomery (2009) salienta que a carta p não é um modelo universal, ou seja, não é aplicável para todos os contextos, porém possui grande aplicabilidade dentro de seus limites de atuação. Neste caso, trata-se de uma abordagem focada no monitoramento de características que se adequam à distribuição Binomial, ou seja, na avaliação de pontos que se enquadram ou não em determinada característica avaliada (CHUKHROVA e JOHANNSEN, 2019). Segundo Hart, Hart e Schmaltz (2007), as abordagens convencionais das cartas p adotam os limites de controle distantes em 3 desvios-padrão da linha central, sendo tal valor obtido a partir da aproximação da distribuição Binomial por uma Normal. Dessa forma, os limites de controle da carta p , LSC e LIC, são obtidos por meio das equações 2 e 3,

respectivamente. Além disso, Linha Central (LC) é a proporção da característica avaliada no conjunto de dados (ABBAS *et al.*, 2019).

$$LSC = p + 3 \sqrt{\frac{p(1-p)}{n}} \quad (2)$$

$$LIC = p - 3 \sqrt{\frac{p(1-p)}{n}} \quad (3)$$

Em que p é a proporção da característica avaliada e n é o tamanho da amostra adotada.

Nota-se que os limites LSC e LIC são dependentes do tamanho da amostra ' n ' e da proporção ' p '. Logo, tratam-se de parâmetros chave da carta e, conseqüentemente, merecem atenção especial durante a fase I da do estudo. Com relação ao valor da proporção ' p ', este é dependente do próprio processo que se deseja monitorar e assume-se que se trata de um valor conhecido. Porém, para casos onde a proporção p é desconhecida, esta deve ser estimada a partir das observações do processo (MONTGOMERY, 2009). Neste caso, o autor destaca que o procedimento usual é selecionar amostras preliminares (entre 20 e 25) visando obter a proporção média dessas amostras. Por fim, a proporção estimada ' \bar{p} ' deve ser adotada nas equações 2 e 3 para se obter os limites de controle. Já com relação ao tamanho da amostra, Chukhrova e Johannssen (2019) afirmam que as cartas p são geralmente projetadas com uma amostra de tamanho constante. Além disso, Montgomery (2009) destaca que este valor deve ser grande o suficiente para detectar não conformidades no processo capazes de tira-lo do estado de "sob controle". Neste caso, o autor apresenta a Equação 4 para calcular o valor ' n '.

$$n = \left(\frac{L}{\delta}\right)^2 p(1-p) \quad (4)$$

Em que p é a fração da característica avaliada, L a distância dos limites de controle da linha central e δ é a magnitude do deslocamento do processo. Neste caso, considerando que \hat{p}_0 é o valor estimado em controle de p e \hat{p}_1 é o fora de controle, então temos que $\hat{p}_1 = \hat{p}_0 + \delta$.

Por fim, vale destacar a importância de se avaliar o desempenho da carta de controle visando garantir sua eficácia quanto ao monitoramento dos sistemas. Neste caso, Aebtarm e Bouguila (2011) sugerem alguns indicadores de avaliação do gráfico de atributos, os quais podem ser

adotados para a carta p : (i) Qualidade da plotagem; (ii) *Average Run Length* (ARL) ou número médio de amostras até o sinal (NMA), em português; e (iii) *Loss Function*, ou função de perda em português. Tais indicadores referem-se à capacidade da carta em apresentar claramente o comportamento do sistema monitorado, à taxa de alarmes falsos da carta e ao custo associado à presença desses alarmes falsos, respectivamente. Neste contexto, o mais adotado é o ARL, o qual pode ser obtido por meio da equação 5. Os alarmes falsos a que se referem tais indicadores ocorrem quando o processo está sob controle, porém a carta sinaliza-o como fora de controle. Neste caso, a probabilidade de tal ocorrência compõe o chamado “Risco α ”. Além desse risco, tem-se o chamado “Risco β ”, o qual refere-se à situação em que a carta deixa de sinalizar quando processo sai fora de controle (COSTA, EPPRECHT e CARPINETTI, 2018).

$$ARL = \frac{1}{pr} \quad (5)$$

Em que pr é a probabilidade de um ponto estar fora de controle (considerando a distribuição Binomial).

3. MÉTODO DE PESQUISA

Este capítulo apresenta o método de pesquisa adotado no presente trabalho. Primeiramente será apresentada a classificação da pesquisa, com destaque quanto à sua natureza, objetivo, abordagem e método. Em seguida serão descritos, de maneira mais detalhada, os métodos de pesquisa adotados, a Experimentação e Modelagem e Simulação, com ênfase em suas características e etapas principais. Por fim, são apresentadas as etapas a serem seguidas neste trabalho

3.1. Classificação da pesquisa

Segundo Prodanov e Freitas (2013), pode-se definir a pesquisa como sendo um estudo planejado acerca de um problema ou situação. Vale ressaltar que o método de abordagem deste estudo é o que caracteriza o seu aspecto científico. Neste caso, a classificação da pesquisa é imprescindível para o direcionamento das atividades a serem executadas e é diretamente ligada ao ponto de vista adotado. Dessa forma, pode-se classificar a pesquisa quanto à sua natureza, seu objetivo, sua abordagem e, por fim, quanto ao método adotado (MIGUEL *et al.*, 2018; PRODANOV e FREITAS, 2013). A seguir são apresentadas as principais classificações da presente pesquisa.

3.1.1. Quanto à natureza

Segundo Prodanov e Freitas (2013), uma pesquisa pode ser classificada como: (i) **Básica**: cujo objetivo é gerar conhecimentos científicos sem aplicação prevista; e (ii) **Aplicada**: quando objetiva-se a aplicação prática em problemas específicos a partir dos conhecimentos gerados. Dessa forma, com base em tais considerações, pode-se classificar a presente tese como uma pesquisa **Aplicada**. Neste caso, trata-se de um trabalho focado na proposição de uma solução com viés prático, testado em objetos de estudo teóricos e práticos e com potencial de ampla adoção por parte dos processos produtivos.

3.1.2. Quanto ao objetivo

Já com relação ao seu objetivo, Gil (2002) relata que uma pesquisa pode ser classificada como: (i) **Exploratória**: na qual objetiva-se compreender e formular hipóteses acerca de determinados fenômenos; (ii) **Descritiva**: quando o objetivo principal é descrever características de determinados fenômenos; e (iii) **Explicativa**: quando o foco é identificar fatores que interferem em determinados fenômenos. Além disso, a pesquisa de caráter explicativo está vinculada a sistemas e processos reais e relacionada, em sua maioria, com métodos experimentais. Desta forma, a presente pesquisa pode ser classificada como **Exploratória** e **Explicativa**. Em primeiro lugar, considera-se a pesquisa como Exploratória devido ao fato de que um dos objetivos deste trabalho é explorar o estado da arte sobre o tema abordado. Por outro lado, entende-se que o trabalho é majoritariamente Explicativo, já que serão analisados fenômenos de processos reais visando obter relações entre estes e seus parâmetros, bem como serão propostas análises periódicas do mesmo a fim de se garantir seu correto funcionamento.

3.1.3. Quanto à abordagem

Miguel *et al.* (2018) dão destaque para dois tipos de abordagens de pesquisa: (i) **Qualitativa** e (ii) **Quantitativa**. A abordagem Qualitativa baseia-se na interpretação de um fenômeno visando descrever, decodificar e entender suas variáveis a partir de práticas com pouca ou nenhuma quantificação. Por outro lado, na abordagem Quantitativa parte-se de teorias para posteriormente seguir para formulação de hipóteses, coleta dados a partir de observações, análise dos dados e, por fim, geração de resultados. Neste contexto, pode-se classificar a presente tese como sendo uma pesquisa **Qualitativa** e **Quantitativa**. A classificação

Qualitativa se deve ao fato de que um dos objetivos específicos deste trabalho é avaliar a literatura quanto ao estado da arte referente ao tema abordado. Além disso, considera-se a pesquisa como majoritariamente Quantitativa, uma vez que se pretende analisar sistemas reais de maneira a mensurar e analisar seus parâmetros de operação. Neste caso, pode-se destacar etapas de coleta de dados dos sistemas físicos, execução de experimentos e análises e, posteriormente, interação com tais sistemas de forma a interferir no ambiente real baseando-se nos resultados obtidos.

3.1.4. Quanto ao método

Por fim, é possível classificar a pesquisa por meio do método utilizado para seu desenvolvimento. Neste caso, dentre os principais métodos dentro da abordagem quantitativa, Miguel *et al.* (2018) destacam: (i) Pesquisa de Avaliação (*survey*): a qual consiste na coleta e análise de informações acerca de um determinado fenômeno/característica a partir da aplicação de questionários ou outro instrumento a um grupo pré-definido; (ii) Modelagem e Simulação: a qual se baseia em abstrações de determinados fenômenos ou sistemas a partir de modelos capazes de permitir a identificação de problemas, coleta de informações, formulação de hipóteses e auxílio para a tomada de decisões; e (iii) Experimentação: onde testa-se o relacionamento entre variáveis de determinado fenômeno ou sistema de forma a se manipular certos comportamentos e obter relações de causa e efeito importantes para a compreensão dos mesmos. Dessa forma, a presente pesquisa irá basear-se em dois métodos, **Experimentação e Modelagem e Simulação**. Neste caso, considerando os modelos de simulação como GDs adotados neste trabalho, estes foram construídos por meio do método de Modelagem e Simulação. Além disso, toda as etapas de desenvolvimento e melhoria da ferramenta de monitoramento proposta, cujo objetivo é garantir a validade desses modelos durante sua operação, serão baseadas em experimentos. As principais etapas referentes a cada um dos métodos serão apresentadas na seção a seguir. A Figura 3.1 ilustra a classificação da presente pesquisa.

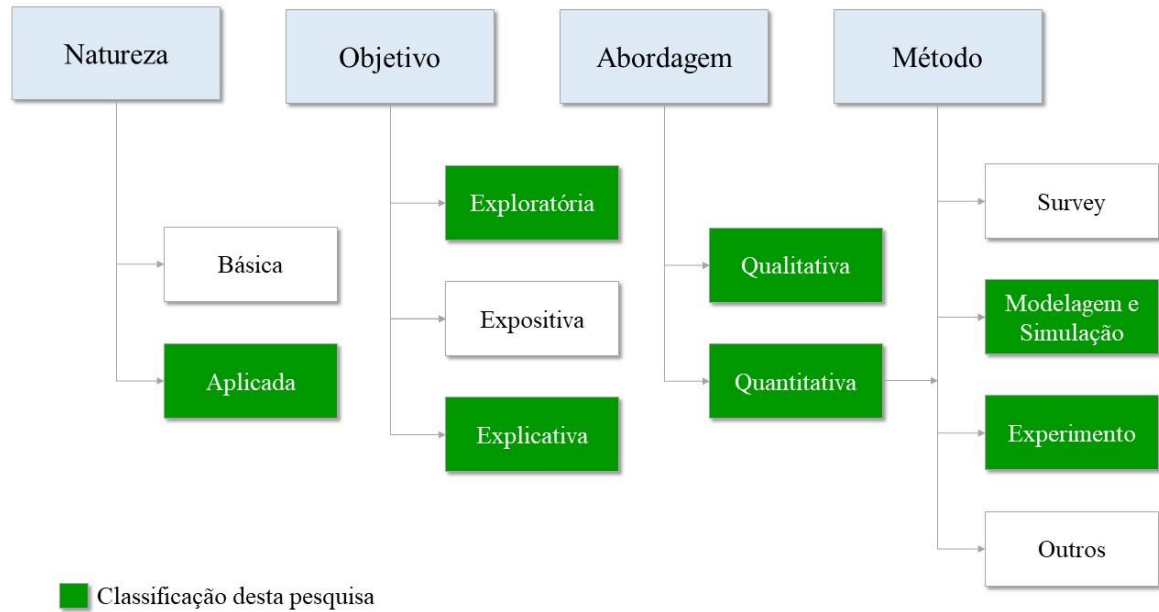


Figura 3.1– Classificação da pesquisa

Fonte: adaptado de Prodanov e Freitas (2013) e Miguel *et al.* (2018)

3.2. Método

3.2.1. Modelagem e Simulação

Dentro do método de Modelagem e Simulação, pode-se considerar a presente pesquisa como **Empírica**. Neste caso, há a busca por correspondência entre o modelo de simulação e os sistemas reais visando aprimorar o processo de tomada de decisão (BERTRAND e FRANSOO, 2002; MIGUEL *et al.*, 2018). A literatura apresenta diversos métodos e *frameworks* contendo as principais etapas consideradas no desenvolvimento de projetos de simulação neste contexto (MONTEVECHI *et al.*, 2015). Além disso, considerando sua utilização como GD, são necessárias algumas outras etapas chave visando integrar o modelo de simulação com os sistemas físicos (SANTOS *et al.*, 2022a). Dessa forma, para a construção dos modelos de simulação como GDs utilizados nesta tese, optou-se pela utilização do método proposto por Montevechi *et al.* (2020) e Santos *et al.* (2022a). O método proposto é dividido em três grandes fases, chamadas de “Modelagem Conceitual”, “Modelagem Computacional” e “Modelagem Operacional”.

A fase de Modelagem Conceitual contempla as seguintes atividades: (1) Definição dos objetivos e do sistema, (2) Construção do modelo conceitual, (3) Validação do modelo conceitual, (4) Documentação do modelo conceitual e, por fim, (5) Modelagem dos dados de entrada do modelo. Como resultado desta fase, têm-se o Modelo Conceitual do objeto de estudo.

Já a segunda fase, Modelagem Computacional, apresenta as seguintes etapas: (6) Construção do modelo computacional, (7) Verificação do modelo computacional e (8) Definição dos dados de atualização do modelo e das respostas desejadas, (9) Estruturação da interface com o processo real e (10) Validação do modelo computacional. As etapas 8 e 9 são cruciais para a integração do modelo com os sistemas físicos, característica chave do GD. Além disso, a última etapa da fase, validação, garante a validade inicial do modelo, chamado agora de modelo operacional. Finalmente, a última fase, Modelagem Operacional, é composta pelas etapas: (11) Definição dos cenários a serem testados, (12) Execução periódica dos cenários e, por fim, (13) Análise e tomada de decisão. As etapas finais garantem a utilização periódica do modelo para tomada de decisões. Tal método é ilustrado pela Figura 3.2.

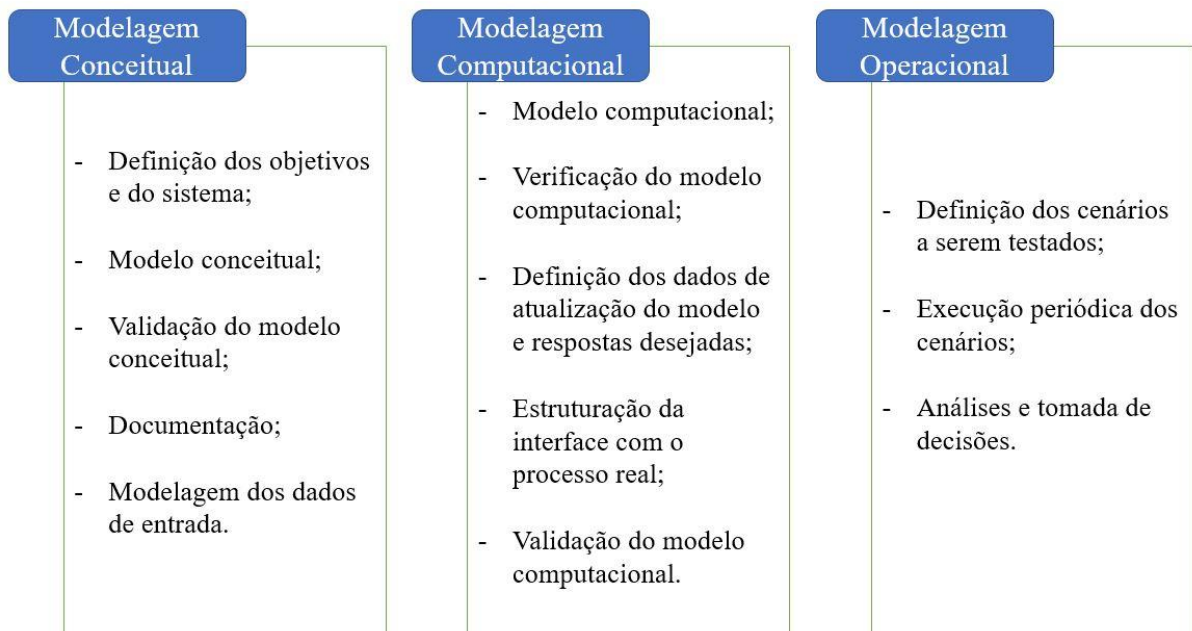


Figura 3.2 – Etapas seguidas na construção dos modelos de simulação

Fonte: adaptado de Montevechi *et al.* (2020) e Santos *et al.* (2022a)

3.2.2. Experimentação

De acordo com Lenhard (2007), a cooperação entre a Modelagem e Simulação e a Experimentação cria uma poderosa fusão metodológica. Neste caso, nota-se que a adoção de ambas representa uma integração sinérgica na busca por soluções e desenvolvimentos. A Experimentação pode ser entendida como um método no qual o pesquisador testa a validade e eficiência de determinadas respostas ou soluções a partir da execução de experimentos, os quais baseiam-se na manipulação planejada de variáveis e fatores controláveis (JUNIOR *et al.*, 2005;

KIDDER, 2004). Além disso, Lazar, Feng e Hochheizer (2017) relatam que o experimento precisa ser planejado e executado de forma sistemática e, neste caso, os autores sugerem cinco etapas principais para tal: (1) Definição do Contexto: na qual o pesquisador define os objetivos, características e escopo do problema a ser abordado; (2) Planejamento: onde são planejados os procedimentos e análises a serem realizados visando atingir os objetivos previstos inicialmente; (3) Execução: onde as atividades planejadas são realizadas, ou seja, trata-se de uma etapa-chave do experimento; (4) Interpretação: na qual o pesquisador avalia os resultados obtidos anteriormente e realiza as inferências com base em seus conhecimentos e com apoio de ferramentas de análise; e (5) Apresentação: onde os resultados finais são compilados e relatados. A Figura 3.3 resume as principais etapas do experimento.



Figura 3.3 – Etapas seguidas na execução dos experimentos

Fonte: adaptado de Lazar, Feng e Hochheiser (2017)

3.3. Etapas da pesquisa

Tendo em vista as discussões e conceitos apresentados até aqui, é importante destacar as etapas consideradas no desenvolvimento desta tese. Neste caso, trata-se de uma base para futuras replicações além de auxiliar no entendimento de toda a estrutura da pesquisa. A primeira etapa da pesquisa refere-se ao embasamento teórico acerca dos temas a serem abordados na tese. Neste caso, além de uma pesquisa exploratória sobre os principais tópicos explorados, optou-se também pela realização de uma RSL visando explorar o estado da arte sobre o uso da simulação como GD. A segunda etapa refere-se ao planejamento da pesquisa, o qual baseia-se nos métodos adotados, tal como descrito no capítulo 3. O capítulo 4 é dedicado ao desenvolvimento prático da pesquisa. Neste caso, primeiramente são apresentados os resultados da RSL que serviu de base para os desenvolvimentos aqui apresentados. Em seguida, é apresentada a abordagem proposta, a qual visa o monitoramento periódico e sistemático dos modelos de simulação utilizados como GD. A seguir, a avaliação da abordagem proposta é apresentada a partir dos resultados experimentais, sendo adotados objetos de estudo teóricos e práticos visando avaliar a performance e eficiência da proposta. Por fim, discussões e conclusões acerca do desenvolvimento são apresentados, de forma a fornecer ao leitor uma

ampla e completa visão sobre o trabalho desenvolvido. O quadro 3.1 resume as principais etapas consideradas nesta pesquisa.

Quadro 3.1 – Etapas consideradas para o desenvolvimento da pesquisa

Etapa	Objetivo	Método de pesquisa adotado
Embasamento teórico	Construção do conhecimento acerca dos tópicos abordados	Revisão da literatura e RSL
Planejamento da pesquisa	Definição dos métodos e etapas	–
Proposição da abordagem	Criação de uma ferramenta de monitoramento dos modelos	Experimentação
Avaliação da abordagem	Teste da abordagem a partir de objetos de estudo teóricos e práticos	Experimentação e Modelagem e Simulação
Discussões e conclusões	Apresentação dos resultados da pesquisa	–

Fonte: elaborado pelo próprio autor

4. APLICAÇÃO

Neste capítulo serão apresentados, primeiramente, os resultados da RSL desenvolvida. Dessa forma, é possível compreender o estado da arte acerca do uso da simulação como GD, bem como explorar as oportunidades relacionadas à área, com foco na validade dos modelos durante sua operação no apoio à decisão. Em seguida, é apresentada a abordagem proposta para o monitoramento de modelos de simulação como GDs. Neste caso, serão descritas as características principais da abordagem, bem como suas etapas de implementação. Por fim, visando analisar a aplicabilidade da ferramenta de monitoramento proposta, optou-se pela sua implementação em objetos de estudo teóricos e reais de forma a se avaliá-la sob diferentes perspectivas de uso. Conforme destacado anteriormente, esta tese não é focada nas etapas relacionadas à construção dos modelos de GD, sendo adotados para tal, os métodos descritos no capítulo 3 deste trabalho.

4.1. Estado da arte acerca do uso da simulação como Gêmeo

Digital: uma Revisão Sistemática da Literatura

De forma a explorar o estado da arte quanto à utilização da simulação como GD de processos produtivos, foi desenvolvida uma revisão sistemática da literatura. Enquanto uma revisão de

literatura (RL) permite o desenvolvimento de uma pesquisa exploratória sem necessariamente seguir padrões metodológicos rigorosos, a Revisão Sistemática da Literatura (RSL) aborda a literatura de maneira estruturada baseando-se em etapas bem definidas para responder certas perguntas científicas (OLIVEIRA, LIMA e MONTEVECHI, 2016; TRIGUEIRO *et al.*, 2019). Tranfield, Denyer e Smart (2003) ressaltam que a RSL é uma atividade científica fundamental, a qual baseia-se em duas frentes: (i) varredura científica sobre as principais obras de determinada área e (ii) procedimentos estatísticos visando sintetizar as descobertas e dar credibilidade aos resultados. Além disso, para o desenvolvimento da RSL, optou-se por estruturar a pesquisa de acordo com os passos sugeridos por Oliveira, Lima e Montevechi (2016), que são: (i) Planejamento (definição dos objetivos e questões de pesquisa); (ii) Pesquisa/triagem (varredura da literatura); (iii) Análise (análises dos resultados e procedimentos estatísticos); e (iv) Apresentação (descrição dos resultados e principais conclusões). Dessa forma, o desenvolvimento da RSL é apresentado nos tópicos seguintes.

4.1.1. Planejamento

De forma a definir os objetivos e questões-chave da pesquisa, o primeiro passo foi realizar uma pesquisa exploratória sobre o tema abordado nesta pesquisa. Neste caso, utilizou-se da base de dados Scopus® que, de acordo com Scheidegger *et al.* (2018) é uma das principais bases de dados multidisciplinares disponíveis. Esta primeira busca por artigos contemplou as palavras-chave “*Simulation*” e “*Digital Twin*”. Neste primeiro momento, não foi especificado o tipo de simulação, visto que se trata apenas de uma primeira análise exploratória. Além disso, focou-se em trabalhos voltados para a tomada de decisões em processos produtivos, publicados em jornais científicos e anais de congressos e avaliadas por pares. Além dos artigos encontrados nesta primeira etapa, outros trabalhos referenciados por eles foram lidos oportunamente. A partir desta etapa inicial, os pesquisadores identificaram as principais nomenclaturas utilizadas para se referir à utilização da simulação como GD, além de certas questões ainda não definidas sobre tal abordagem, as quais foram base para a formulação das questões de pesquisa (QPs). Quanto à área de atuação, nota-se trabalhos focados na produção tanto de bens como serviços (LOPES *et al.*, 2019; STERINGER *et al.*, 2019). Já com relação a métodos para desenvolvimento dos modelos, observou-se apenas a presença de *frameworks* específicos, sem a existência de métodos mais gerais (LU *et al.*, 2019; VIJAYAKUMAR *et al.*, 2019). Tantos *softwares* de simulação comerciais como algoritmos personalizados foram identificados (LU *et al.*, 2019; TERKAJ *et al.*, 2019). Quanto à avaliação da validade do modelo de simulação, foi

observado apenas a validação tradicional do modelo, durante sua fase de construção. Já com relação à conexão do modelo com os sistemas físicos, observou-se conexões diretas e também a adoção de sistemas intermediários (KARAKRA *et al.*, 2018; MIETH, MEYER e HENKE, 2019). Além disso, há casos onde os GDs são autônomos e não autônomos (DONHAUSER *et al.*, 2018; EYRE, SCOTT e FREEMAN, 2018). Por fim, há abordagens em tempo real e quase real (BOTTANI, MURINO e VESPOLI, 2017; ZÖRRER *et al.*, 2019).

Baseando-se nas observações mencionadas anteriormente, nota-se que há ainda uma lacuna de trabalhos teóricos que abordem a utilização da simulação como GD. Há diversas aplicações e características as quais devem ser analisadas, comparadas e classificadas visando fornecer uma base teórica sólida sobre o tema à pesquisadores e profissionais da área. A Figura 4.1 ilustra as principais lacunas relacionadas ao tema desta RSL.



Figura 4.1 – Lacunas associadas ao uso da simulação como Gêmeo Digital

Fonte: elaborado pelo próprio autor

Tendo em vista as considerações acima, os objetivos da RSL são:

- (a) Desenvolver uma extensa varredura na literatura sobre aplicações da simulação como GD;
- (b) Analisar cada trabalho sobre o ponto de vista das lacunas identificadas;
- (c) Criar uma base teórica sólida acerca da utilização da simulação neste contexto;
- (d) Identificar perspectivas futuras relacionadas ao tema.

Uma vez definidos os objetivos, as QPs foram formuladas de acordo com o escopo do estudo:

QP 1: Quais as áreas de atuação e os objetivos de decisão associados ao uso da simulação como Gêmeo Digital?

QP 2: Quais as plataformas utilizadas para modelagem e como é realizada a sua conexão com os sistemas físicos?

QP 3: Qual o horizonte de tempo considerado para a atualização do modelo de simulação frente às mudanças físicas?

QP 4: Qual o nível de autonomia dos modelos de simulação como Gêmeos Digitais?

QP 5: Há métodos para construção e avaliação periódica da validação de modelos dessa natureza?

QP 6: Quais as principais vantagens, desafios e oportunidades atreladas a esta abordagem?

4.1.2. Pesquisa/triagem

O primeiro passo desta etapa foi a seleção das bases de dados a serem utilizadas na varredura da literatura. Neste caso, foram selecionadas cinco (5) bases: Scopus®, Web of Science®, Scielo®, IEEE Xplore® e Science Direct®. Em seguida, as palavras-chave de pesquisa foram definidas. Para obter maior precisão, o termo “*Simulation*” foi estratificado em seus principais tipos considerados em processos produtivos, conforme já mencionado: “*Discrete-event Simulation*” e “*Agent-based Simulation*”. Além disso, foram incluídos os termos que se referem a utilização da simulação como GD, conforme já mencionado. São eles: “*Digital Twin*”, “*Cyber-physical Simulation*”, “*Real-time Simulation*”, “*Near Real-time Simulation*”, “*Symbiotic Simulation*”, “*Online Simulation*”, “*Data-driven Simulation*” e “*Semi-physical Simulation*”. As lógicas booleanas (AND / OR) foram utilizados visando obter todas as combinações possíveis entre os termos. Ao todo, foram realizadas 80 pesquisas, considerando todas as bases de dados escolhidas. Além disso, em cada pesquisa foram incluídos os seguintes critérios para se considerar o artigo na RSL: (i) os termos pesquisados devem estar presentes no título, resumo ou palavras-chave; (ii) artigos publicados nos últimos trinta anos, considerando a data final como sendo Março de 2020; (iii) artigos completos publicados em revistas científicas e anais de congressos e que tenham sido avaliados por pares; (iv) somente trabalhos escritos na língua inglesa; (v) artigos de teor prático e com foco na tomada de decisões de operações produtivas (bens e serviços).

Dessa forma, em um primeiro momento foram encontrados 169 artigos que se enquadram nos critérios de pesquisa, já descontando as redundâncias. Após esta primeira seleção de artigos, foi realizada uma etapa de triagem com todos os trabalhos, onde foram lidos os resumos visando

identificar aqueles que estão alinhados aos objetivos da RSL. Nesta etapa, foram excluídos artigos que não se adequavam aos critérios de pesquisa anteriormente definidos. Após a triagem, 75 artigos foram considerados para leitura completa. A Figura 4.2 resume os procedimentos realizados nas duas primeiras etapas da RSL.

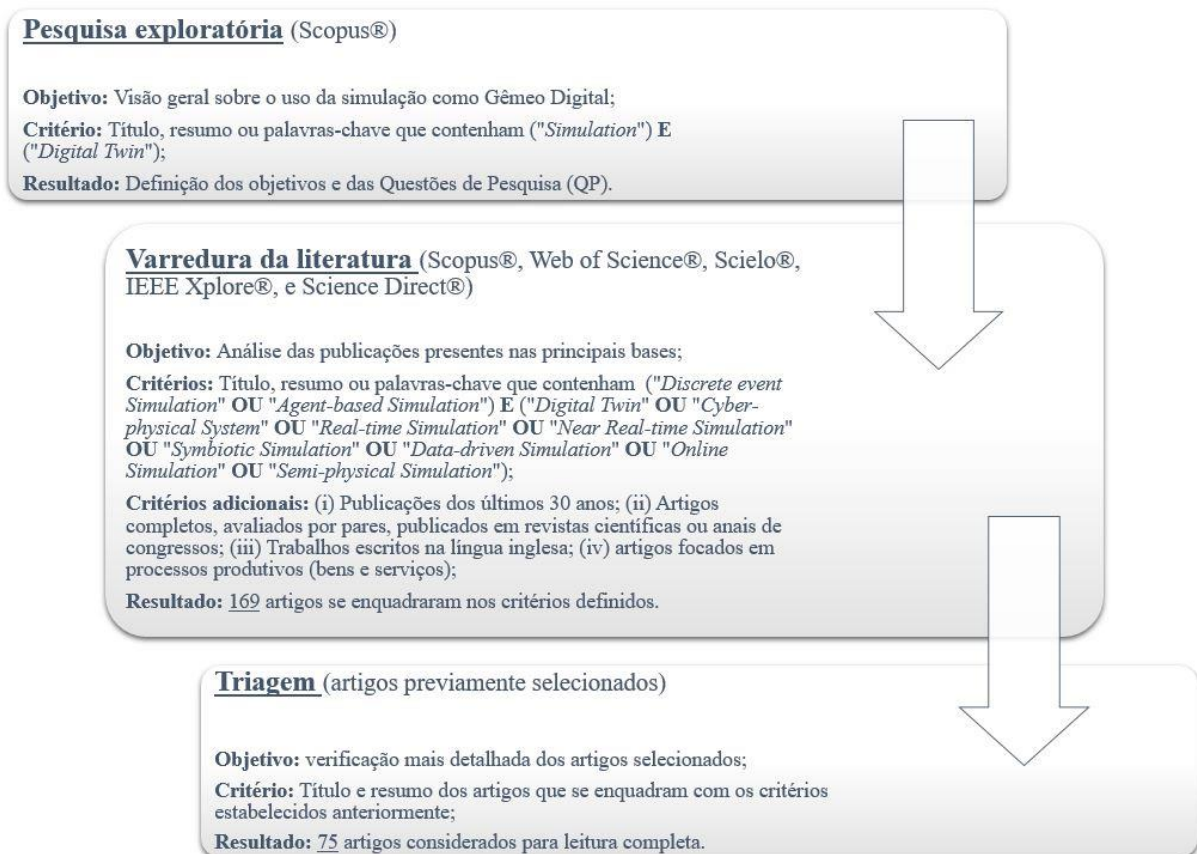


Figura 4.2 – Procedimentos realizados nas etapas de planejamento e pesquisa/triagem

Fonte: elaborado pelo próprio autor

4.1.3. Análise e apresentação

As análises e sínteses dos resultados foram realizadas por meio de uma planilha MS Excel®. Dessa forma, foi possível compilar as informações de todos os artigos analisados. As análises foram baseadas em cada uma das QPs e demonstram as melhores práticas e principais perspectivas relacionadas à utilização da simulação como GD. A apresentação dos resultados será descrita nos tópicos a seguir e, neste caso, será apresentado primeiramente uma visão geral sobre os trabalhos analisados, seguidas das respostas relacionadas a cada uma das RQs definidas anteriormente.

I. Visão Geral dos trabalhos

De uma maneira geral, observa-se que há certo equilíbrio entre o número de publicações em revistas e anais, conforme ilustra a Figura 4.3. Os 75 artigos analisados foram publicados em 25 revistas científicas e 24 anais de congressos. Além disso, fica claro que há uma tendência de crescimento das publicações ao longo do período analisado. Neste caso, o destaque é para os últimos cinco anos, período no qual concentrou-se cerca de 58,7% do total de trabalhos publicados. O crescimento acentuado de publicações recentemente pode ser associado ao campo fértil de pesquisas relacionadas a implementação de GDs (TAO e ZHANG, 2017).

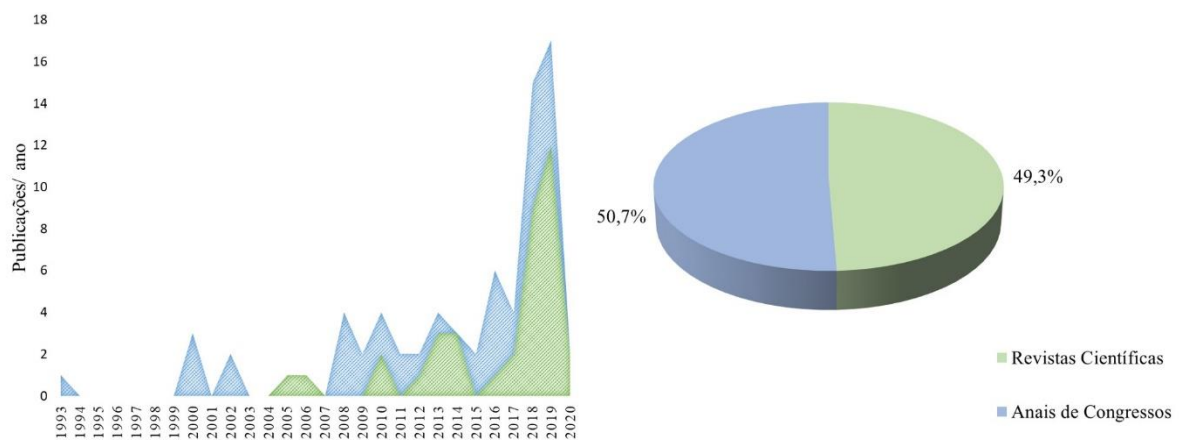


Figura 4.3 – Publicações quanto à simulação como Gêmeo Digital

Fonte: elaborado pelo próprio autor

Além disso, é possível também estratificar as publicações em função da nomenclatura utilizada para se referir a utilização da simulação como GD. Percebe-se na Figura 4.4 que o termo “*Digital Twin*” aparece em 29,3% das publicações, sendo presente a partir de 2016. Uma vez que o termo “*Digital Twin*” foi proposto em 2010, seis anos se passaram até que a simulação pudesse ser vista como uma alternativa para sua concepção. Por outro lado, o primeiro trabalho que remete à ideia da utilização da simulação como GD foi publicado em 1993, o qual adotou a nomenclatura “*Online Simulation*”. Os termos “*Near Real-time*” e “*Real-time*” correspondem à maior parte das publicações (34,7%), enquanto o termo “*Cyber-physical System*” também representa uma quantidade significativa de trabalhos (16%). Nota-se que o termo “*Semi-physical Simulation*” não apareceu em nenhum trabalho e, por fim, nota-se que os termos “*Symbiotic*” e “*Online Simulation*” deixaram de ser utilizados nos últimos anos, fato que pode estar relacionado com a criação do termo “*Digital Twin*” para se referir a modelos de simulação desta natureza.

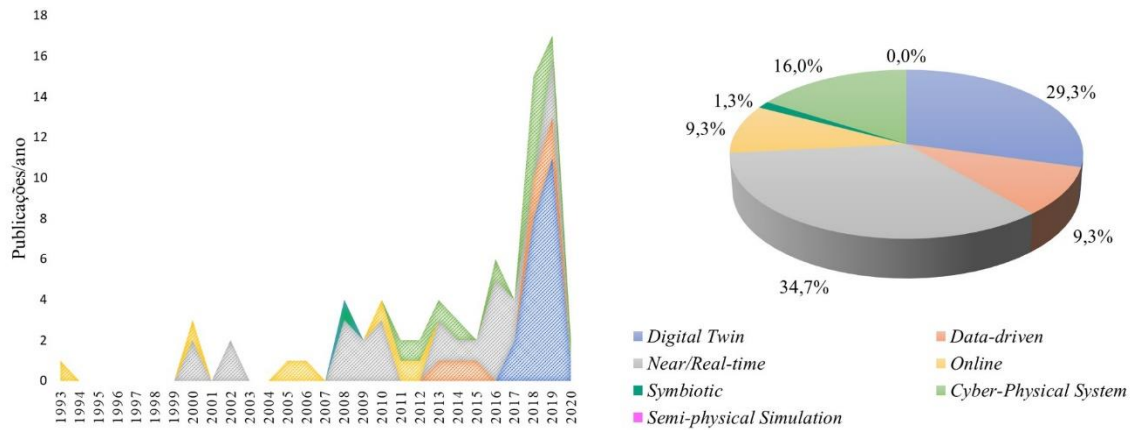


Figura 4.4 – Publicações quanto à simulação como Gêmeo Digital (estratificado)

Fonte: elaborado pelo próprio autor

Finalmente, pode-se avaliar as publicações em função do tipo de simulação utilizado. Observa-se que a SED está presente em 92% das publicações, seguido por 5,3% dos trabalhos que utilizam a SBA e, por fim, 2,7% que utilizam ambos os tipos. Neste último caso, há o que chamamos de simulação híbrida, onde mais de um tipo de simulação é utilizado de maneira integrada (BRAGLIA *et al.*, 2019).

II. Áreas e objetivos quanto ao uso da simulação como Gêmeo Digital

Quanto às áreas de atuação e os objetivos da tomada de decisão por meio da simulação como GD, nota-se que há grande diversidade entre os trabalhos analisados. Primeiramente, com base nos artigos selecionados, dividimos as áreas de atuação em seis principais: (i) Manufatura (corresponde a atividades relacionadas à produção de bens); (ii) Serviços (relacionada a atividades ligadas à serviços, tal como laboratórios, serviços de TI, etc.); (iii) Logística (relacionada a atividades de logística, como roteamento, movimentações de materiais, etc.); (iv) Saúde (relacionadas à operações da área da saúde tal como hospitais, clínicas, etc.); (v) Construção (relacionada a atividades ligadas à construção civil); e (vi) Outras (as quais não se enquadram em nenhuma das demais áreas). Dessa forma, a Figura 4.5 apresenta o percentual de artigos ligados a estas áreas de atuação.

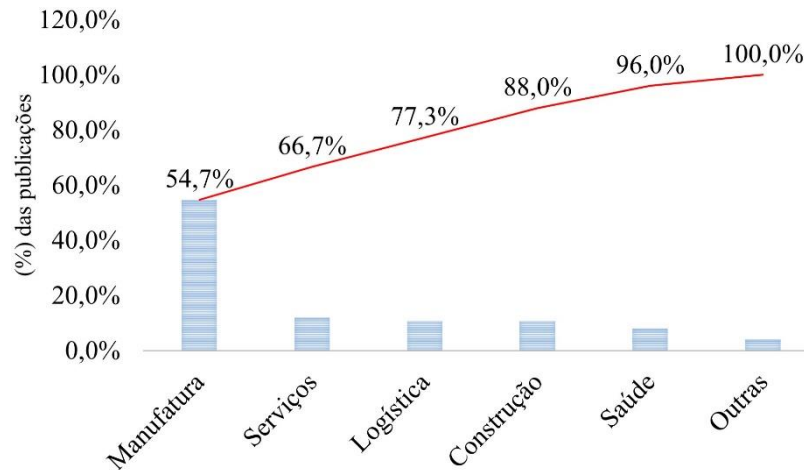


Figura 4.5 – Principais áreas de atuação quanto à simulação como Gêmeo Digital

Fonte: elaborado pelo próprio autor

Observa-se que a área de Manufatura é responsável por cerca de 54,7% das publicações avaliadas. Neste caso, destaca-se aplicações no setor aeronáutico (STERINGER *et al.*, 2019) automotivo (LU *et al.*, 2019) e metal-mecânico (TERKAJ *et al.*, 2019). A área de Serviços representa cerca de 12% dos trabalhos e cobre processos relacionados a atividades laboratoriais (LOPES *et al.*, 2019), setor energético (LIN *et al.*, 2011, 2012; THANOS *et al.*, 2017; WAN *et al.*, 2014) e TI (BRADFORD, SIMMONDS e UNGER, 2000; SIMMONDS, BRADFORD e UNGER, 2014; XIAOBO *et al.*, 2009). A terceira área mais representativa é a Logística (cerca de 10,7% dos artigos), abrangendo diferentes escopos (ASHRAFIAN *et al.*, 2019; RENIERS e MORTEL-FRONCZAK, 2018; VIJAYAKUMAR *et al.*, 2019). Com o mesmo percentual, a área de Construção apresenta propostas para diferentes características de decisão (AKHAVIAN e BEHZADAN, 2014; ELNIMR, FAGIAR e MOHAMED, 2016; LILIS, VAN CUTSEM e KAYAL, 2019; WANG, QIANCHUAN e YIN, 2013). Na Saúde (8% dos artigos), observa-se GDs nas mais diversas áreas hospitalares (HARPER e MUSTAFEE, 2019b; SORMAZ e MALIK, 2018). Por fim, a área Outros (4% dos artigos) é marcada por aplicações específicas, como atividades com drones (KHALEGHI *et al.*, 2013) e operações militares (ZHOU, HUANG e HU, 2008).

Já com relação aos objetivos da tomada de decisão, observa-se também seis principais, de acordo com os artigos avaliados. Desta forma, os principais objetivos encontrados foram: (i) Planejamento da produção (relacionado ao planejamento de processos, *scheduling*, planejamento de paradas e manutenções, etc.); (ii) Avaliação de processos (associado ao monitoramento dos processos por meio de indicadores de performance); (iii) Controle de processos (associado ao controle dos sistemas físicos) (iv) Alocação de recursos (o GD oferece

diretivas relacionadas à alocação de recursos); (v) Roteamento (associado a análises e otimizações de rotas para equipamentos e pessoas); e (vi) Outros (os quais não se enquadram em nenhum dos demais objetivos). A Figura 4.6 apresenta o percentual de artigos ligados a cada um destes objetivos.

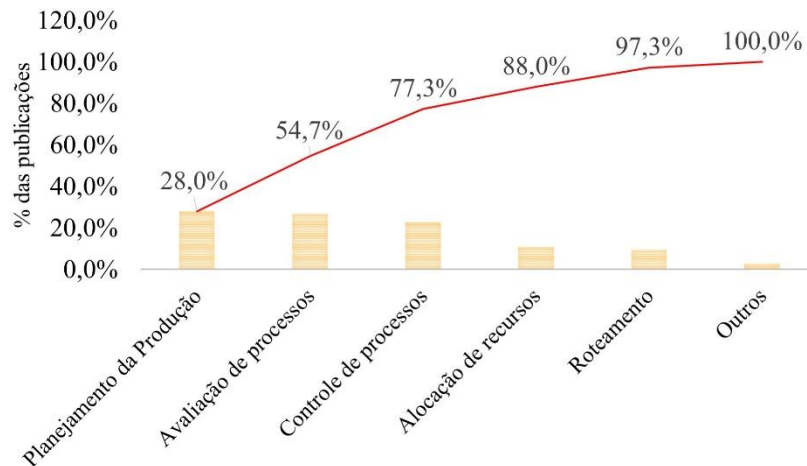


Figura 4.6 – Principais objetivos quanto à simulação como Gêmeo Digital

Fonte: elaborado pelo próprio autor

Observa-se que as decisões relacionadas ao Planejamento da Produção e Avaliação de Processos são responsáveis por mais da metade das publicações (cerca de 54,7%). Considerando o Planejamento da Produção, nota-se aplicações em processos específicos e também abordagens generalistas envolvendo mais de um processo (CHINNATHAI *et al.*, 2018; PRAJAPAT *et al.*, 2019; SHIRAZI, MAHDAVI e SOLIMANPUR, 2010; STERINGER *et al.*, 2019; ZUPAN, JANEZ e HERAKOVIČ, 2018). Quanto à Avaliação de Processo, os principais focos estão relacionados à avaliação de desempenho (CHOI e KANG, 2018; EYR, SCOTT e FREEMAN, 2018; LU *et al.*, 2019; SALAMA e ELTAWIL, 2018; VIJAYAKUMAR *et al.*, 2019), mas há abordagens vinculadas a objetivos comerciais (TERKAJ *et al.*, 2019). Considerando Controle de Processos, grande parte dos trabalhos se refere a sistemas automatizados (SHANG e WAINER, 2008; SICKEL e LEE, 2009), porém há também controles indiretos em processos manuais (WANG *et al.*, 2013). No caso da Alocação de Recursos, nota-se artigos focados em recursos humanos e também materiais (AUGUSTO, MURGIER e VIALON, 2018; LOPES *et al.*, 2019; THANOS *et al.*, 2017). Por fim, em relação ao Roteamento, há casos de otimização de rotas para veículos automatizados e não automatizados (ASHRAFIAN *et al.*, 2019; BARCIK, MÖLLER e VAKILZADIAN, 2016; BOTTANI, MURINO e VESPOLI, 2017; ELNIMR, FAGIAR e MOHAMED, 2016). A categoria Outros

é representada por objetivos específicos, como decisões relacionadas a processos militares (ZHOU, HUANG e HU, 2008).

III. Modelagem e conexão dos modelos com os sistemas físicos

Considerando as plataformas de modelagem, diversas práticas foram observadas. Cerca de 54,7% dos trabalhos avaliados utilizaram 16 *softwares* e pacotes comerciais diferentes. Nesse caso, os 10 mais utilizados correspondem a 85,4% das publicações e incluem: Tecnomatix® (24,4%), Arena® (14,6%), AnyLogic® (9,8%), FlexSim® (7,3%), Quest® (4,9%), Simio® (4,9%), Simul8® (4,9%), Repast® (4,9%), Symphony® (4,9%) e Network Simulator® (4,9%). Nesse caso, a preferência pelos pacotes comerciais se deve principalmente aos recursos disponíveis, como interface gráfica, fácil construção de modelos e relatórios customizados (LU *et al.*, 2019). Por outro lado, aproximadamente 32% dos artigos relataram o uso de linguagens de programação para desenvolver os algoritmos de simulação. Dentre as linguagens utilizadas, destacamos o Python® (LILIS, VAN CUTSEM e KAYAL, 2019; OLAITAN *et al.*, 2014), Java® (TIACCI, 2020), C++® (MOALLEMI e WAINER, 2010) e Stroboscope® (AKHAVIAN e BEHZADAN, 2018). O uso de linguagens de programação pode resolver algumas dificuldades encontradas em soluções comerciais, como a falta de recursos para determinadas áreas de aplicação e investimentos consideráveis (CHONG e SIVAKUMAR, 2002; ELNIMR, FAGIAR e MOHAMED, 2016; OLAITAN *et al.*, 2014). Por fim, em cerca de 13,3% dos artigos, não foi possível identificar a plataforma utilizada para modelagem.

Já com relação às formas de conexão entre o modelo de simulação e os sistemas físicos, observou-se também diferentes práticas adotadas pelos trabalhos analisados. A conexão entre o modelo e os sistemas físicos é pré-requisito para que se possa implementar um GD por meio da simulação e, o nível de integração entre os ambientes físico e virtual depende das características de cada aplicação. Ashrafian *et al.* (2019) relatam que os dados das operações são a base para a concepção do GD por meio da simulação e, neste caso, dentre as principais fontes de dados encontram-se os sistemas de gestão, bases de dados, sensores e dispositivos IoT (ALAM e SADDIK, 2017). A Figura 4.7 apresenta a proporção de trabalhos que adotam cada tipo de conexão.

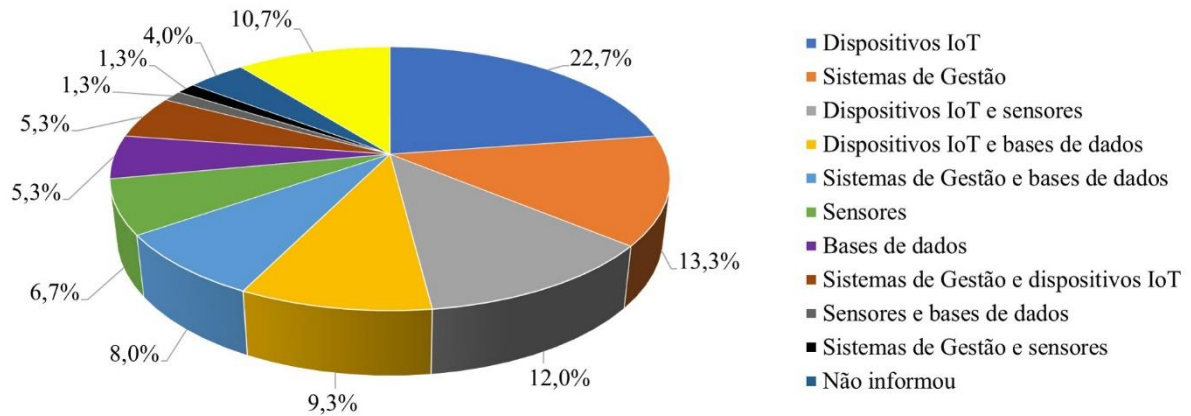


Figura 4.7 – Conexão entre os modelos de simulação e os sistemas físicos

Fonte: elaborado pelo próprio autor

As conexões com dispositivos IoT correspondem a 22,7% das publicações, enquanto 13,3% dos trabalhos adotam apenas a conexão com sistemas de gestão e 6,7% se conectam com sensores. Por fim, cerca de 5,3% adotam apenas a conexão com bancos de dados. Além disso, cerca de 37,3% dos artigos adotaram mais de uma fonte de conexão para obter dados de sistemas reais. Entre as conexões com dispositivos IoT, destacam-se controladores lógico-programáveis (GUPTA e SIVAKUMAR, 2005; SAEZ *et al.*, 2018; SIVAKUMAR e GUPTA, 2006), dispositivos *Bluetooth* e GPS (KITAZAWA *et al.*, 2016; SONG, RAMOS e ARNOLD, 2008), veículos de transporte automatizados (BOTTANI, MURINO e VESPOLI, 2017; KATZ e MANIVANNAN, 1993), entidades com comunicação via radiofrequência (ALTAF *et al.*, 2015; GOODALL, SHARPE e WEST, 2019), entre outros.

Além disso, quando se trata de sistemas de gestão, há uma ênfase nos sistemas de gestão empresariais ou *Enterprise Resource Planning (ERP)* (BERGMANN, STELZER e STRASSBURGER, 2011; MENG *et al.*, 2013; STERINGER *et al.*, 2019). Três estudos (4%) não detalham os tipos de conexão adotados (D'ANGELO; CHONG, 2018; OLAITAN *et al.*, 2014; TIACCI, 2020). Por fim, oito estudos (10,7%) não apresentaram qualquer forma de conexão com os sistemas físicos, assemelhando-se a abordagens tradicionais de simulação (ANTONS e ARLINGHAUS, 2020; ASHRAFIAN *et al.*, 2019; BARCIK, MÖLLER e VAKILZADIAN, 2016; CHINNATHAI *et al.*, 2018; ELNIMR, FAGIAR e MOHAMED, 2016; LOPES *et al.*, 2019; RENIERS e MORTEL-FRONCZAK, 2018; WANG, QIANCHUAN e YIN, 2013).

IV. Horizonte de tempo considerado para a atualização do modelo de simulação

Em se tratando da atualização e sincronismo dos modelos de simulação frente às mudanças físicas, observa-se duas abordagens: atualização em tempo real e atualização em tempo quase real. Dentre os estudos analisados, a grande maioria (cerca de 72%) adota a abordagem em tempo quase real e observa-se aplicações onde o intervalo de tempo entre as atualizações é fixo (BRAGLIA *et al.*, 2019; CASSETTARI *et al.*, 2017; KHALEGHI *et al.*, 2013; STERINGER *et al.*, 2019), e também abordagens onde a atualização do modelo ocorre em diferentes intervalos de tempo, variando de acordo com a necessidade de tomar decisões (HARPER e MUSTAFEE, 2019b; LILIS, CUTSEM e KAYAL, 2016; SIVAKUMAR e CHONG, 2000). Já com relação aos trabalhos que adotam a atualização em tempo real (cerca de 28%), nota-se que os objetivos quase sempre estão relacionados à avaliação e/ou controle do processo (BEREGI, SZALLER e KÁDÁR, 2018; KITAZAWA *et al.*, 2016b; ROSSMANN, SCHLUSE e WASPE, 2008; WAN *et al.*, 2014). Saez *et al.* (2018) revelam que a complexidade dos sistemas físicos, a natureza estocástica dos dados e a automação necessária são fatores que dificultam a adoção de sistemas automatizados e atualização em tempo real, sendo a abordagem em tempo quase real uma boa alternativa para suporte à decisão (ALAM e SADDIK, 2017).

V. Autonomia dos modelos de simulação como Gêmeos Digitais

O grau de autonomia dos modelos de simulação como GDs é outra questão importante e que pode gerar certa confusão quanto classificar ou não um modelo como tal. O grau de autonomia de um GD está relacionado diretamente com o formato de suas respostas, as quais podem exercer comando direto no sistema físico (autônomo) ou apenas sugerir ações (não-autônomos) (SANTOS *et al.*, 2022a). Apesar da adoção de sistemas automatizados ser uma tendência crescente nos últimos anos, há ainda certa limitação na adoção de modelos autônomos. Dentre os estudos analisados, apenas 18,7% demonstraram abordagens autônomas (BEREGI, SZALLER e KÁDÁR, 2018; LEE, RAMAKRISHNAN e WYSK, 2002; MOALLEMI e WAINER, 2010; SICKEL e LEE, 2009; THANOS *et al.*, 2017). Observamos que o percentual de artigos com sistema autônomo não condiz com o percentual de GDs voltados para controle de processos (22,7%), conforme apresentado anteriormente. Isso porque há trabalhos em que o controle dos processos ainda carece de intervenção humana para ser realizado, não sendo consideradas abordagens autônomas.

Outro ponto importante em relação à autonomia dos modelos de simulação diz respeito à segurança das operações. Tao e Zhang (2017) revelam que a segurança relacionada à operação

do GD é um fator extremamente importante e que deve ser considerado. Portanto, automatizar a tomada de decisões sem a interferência humana ainda é um desafio. Esse fato pode ajudar a explicar a grande adesão dos sistemas não autônomos (cerca de 70,7% dos artigos). Nesse caso, as respostas dos modelos de simulação são diretrizes e orientações, cabendo ao tomador de decisão cumpri-las ou não. Considerando abordagens não autônomas, as aplicações estão principalmente relacionadas aos objetivos de Planejamento de Processos, Avaliação de Processos, Alocação de Recursos e Roteamento. Além disso, as respostas do modelo de simulação podem ser formatadas por sistemas intermediários, como *dashboards* e planilhas, fornecendo aos tomadores de decisão diretrizes mais precisas (CASSETTARI *et al.*, 2017; MURPHY *et al.*, 2020; TERKAJ *et al.*, 2019). Porém, há casos em que o tomador de decisão deve interpretar as respostas diretamente do modelo de simulação (LOPES, ALMEIDA e ALMADA-LOBO, 2018). Por fim, cerca de 10,7% dos artigos não relataram o grau de autonomia dos modelos de simulação como GDs.

VI. Construção e avaliação periódica dos modelos como Gêmeos Digitais

Os métodos de construção de modelos de simulação são extremamente importantes e apresentam etapas que orientam o seu desenvolvimento (LAW, 2009). No caso de modelos de simulação tradicionais, Montevechi *et al.* (2015) apresentam os principais métodos para construção de modelos de simulação tradicionais, porém, ao se considerar a simulação como GD, alguns aspectos específicos devem ser levados em consideração. Portanto, novas etapas e considerações devem ser incluídas para adaptar o modelo ao escopo e objetivo do GD, como a integração do modelo com os sistemas físicos e sua atualização periódica. Os trabalhos considerados nesta RSL foram analisados de acordo com o método utilizado para o planejamento e implementação dos modelos de simulação e cerca de 90,7% deles não mencionam o método adotado. Esse percentual inclui trabalhos que apenas sugerem *frameworks* gerais sobre a conexão do modelo com os sistemas físicos, sem abordar as etapas seguidas durante a construção do modelo de simulação (EYRE, SCOTT e FREEMAN, 2018; MAHDAVI, SHIRAZI e SOLIMANPUR, 2010; THANOS *et al.*, 2017). Os demais trabalhos (9,3%) apresentam métodos específicos para o estudo abordado, sem representar etapas genéricas (CHONG e SIVAKUMAR, 2002; GOODALL, SHARPE e WEST, 2019; LU *et al.*, 2019; MENG *et al.*, 2013; MOALLEMI e WAINER, 2010; VIJAYAKUMAR *et al.*, 2019).

Já com relação à avaliação periódica dos GDs, Zhuang, Liu e Xiong (2018) relatam que esta é uma das áreas a serem exploradas na literatura. Tal consideração também vale para o caso de modelos de simulação como GDs, logo, há a necessidade de se estabelecer métodos que visam

garantir a validade desses modelos durante sua operação. Para Meng *et al.* (2013), a validação de modelos de simulação é mais crítica quando utilizada como GD, uma vez que há a possibilidade de certos erros irem se acumulando no decorrer das atualizações do modelo. Neste caso, algumas questões importantes a se considerar são relacionadas à velocidade e frequência dos procedimentos de avaliação do modelo (ANAGNOSTOPOULOS e NIKOLAIDOU, 2003). Dessa forma, os artigos selecionados foram analisados quanto às técnicas de avaliação adotadas. Apenas um trabalho adotou a utilização de procedimentos de validação periódica (CHO *et al.*, 2019). Os autores propõem a execução de testes de hipóteses a cada vez que o modelo é atualizado visando comparar seus resultados ao sistema real e, dessa forma, garantir sua validade. Os demais trabalhos (cerca de 98,7%) não abordam a avaliação periódica dos modelos. Por fim, Onggo *et al.* (2018), apesar de não adotarem rotinas periódicas de avaliação periódica, citam a necessidade e importância de tal prática.

VII. Vantagens, desafios e oportunidades atreladas ao uso da simulação como Gêmeo Digital

Apesar da utilização da simulação como GD estar sendo explorada há anos, ainda há inúmeras oportunidades a serem exploradas tanto em termos práticos quanto teóricos, tendo em vista as dificuldades e desafios ainda associados a esta abordagem. Dessa forma, os artigos analisados nesta RSL foram analisados do ponto de vista das vantagens, dificuldades e oportunidades destacadas pelos autores. O quadro 4.1 resume os principais resultados relacionados à esta análise.

Quadro 4.1 – Vantagens, desafios e oportunidades da simulação como Gêmeo Digital

Característica	Resultados
Vantagens	✓ Permite análises em diversas etapas do ciclo de vida dos processos (STERINGER <i>et al.</i> , 2019);
	✓ Permite a integração com outras ferramentas de análise, tais como otimização e previsão, visando maior eficiência da tomada de decisões (DONHAUSER <i>et al.</i> , 2018; ONGGO <i>et al.</i> , 2018; ZÖRRER <i>et al.</i> , 2019);
	✓ É uma alternativa mais simples, mais barata e mais flexível se comparado às soluções de GD comerciais (DONHAUSER <i>et al.</i> , 2018; GOODALL, SHARPE e WEST, 2019; LU <i>et al.</i> , 2019; TERKAJ <i>et al.</i> , 2019);
	✓ Permite uma fácil modelagem dos sistemas físicos visando a criação de um ambiente virtual semelhante e sincronizado (BRAGLIA <i>et al.</i> , 2019);
	✓ Permite uma integração relativamente fácil do modelo de simulação com os sistemas físicos (VIJAYAKUMAR <i>et al.</i> , 2019);
	✓ Permite fácil manipulação do modelo por parte do tomador de decisões, sem a necessidade de especialistas (GRUBE; MALIK; BILBERG, 2019; SORMAZ; MALIK, 2018);

	<p>✓ Proporciona uma revolução na utilização da simulação no apoio à tomada de decisões, com modelos cada vez mais precisos (AKHAVIAN e BEHZADAN, 2018; BEREGI, SZALLER e KÁDÁR, 2018; MIETH, MEYER e HENKE, 2019).</p> <hr/> <p>✓ Apesar da evolução dos <i>softwares</i> de simulação, muitas vezes ainda são necessários sistemas auxiliares para compor a tomada de decisões (STERINGER <i>et al.</i>, 2019);</p> <p>✓ As aplicações ainda são frequentemente associadas às grandes empresas, devido principalmente à estrutura necessária para sua implementação (LU <i>et al.</i>, 2019; OLAITAN <i>et al.</i>, 2014; TERKAJ <i>et al.</i>, 2019);</p> <p>✓ Garantir a qualidade dos dados dos processos ainda é um desafio, fato que impacta diretamente na eficiência da tomada de decisões (MIETH, MEYER e HENKE, 2019; MURPHY <i>et al.</i>, 2020);</p> <p>✓ A velocidade e tempo de processamento dos modelos pode ser um fator crítico em alguns processos (BEREGI, SZALLER e KÁDÁR, 2018; MENG <i>et al.</i>, 2013; THANOS <i>et al.</i>, 2017);</p> <p>✓ A integração entre <i>hardwares</i> e <i>softwares</i> precisa ser aprimorada visando melhor comunicação entre os componentes (MURPHY <i>et al.</i>, 2020; ONGGO <i>et al.</i>, 2018);</p> <p>✓ Há ainda certa limitação do uso da simulação em determinadas áreas de atuação, uma vez que há certa defasagem de recursos já desenvolvidos nas plataformas de simulação, específicos destas áreas (ELNIMR, FAGIAR e MOHAMED, 2016);</p> <p>✓ A segurança relacionada aos modelos é ainda um fator limitante, fato que limita sua autonomia e larga utilização (KARAKRA <i>et al.</i>, 2018);</p> <p>✓ Garantir a validade do modelo durante sua operação é um desafio frente ao carácter dinâmico dos sistemas físicos (HARPER e MUSTAFEE, 2019);</p> <p>✓ A necessidade de alto poder computacional também pode ser um fator crítico em determinadas aplicações (SCHOLL <i>et al.</i>, 2010; ZÖRRER <i>et al.</i>, 2019).</p>
Oportunidades	<hr/> <p>✓ Ampliar o escopo de atuação da simulação como GD, focando em tomada de decisões de nível gerencial (ZÖRRER <i>et al.</i>, 2019);</p> <p>✓ Adotar recursos de realidade virtual e realidade aumentada integrados aos modelos de simulação visando melhores experiências ao tomador de decisões (ZÖRRER <i>et al.</i>, 2019);</p> <p>✓ Simplificar, sistematizar e aprimorar a conexão do modelo com os sistemas físicos (BRAGLIA <i>et al.</i>, 2019; DONHAUSER <i>et al.</i>, 2018; MURPHY <i>et al.</i>, 2020);</p> <p>✓ Desenvolver ferramentas e <i>softwares</i> de modelagem com abordagem genérica, capazes de suportar bibliotecas relacionadas a escopos não convencionais (TIACCI, 2020);</p> <p>✓ Utilizar técnicas de Inteligência Artificial visando garantir o correto funcionamento e aprimoramento dos modelos (EYRE, SCOTT e FREEMAN, 2018; GOODALL, SHARPE e WEST, 2019);</p>

✓ Desenvolver procedimentos e métodos de avaliação periódica dos modelos de simulação (HARPER e MUSTAFEE, 2019; ONGGO *et al.*, 2018; SCHOLL *et al.*, 2012);

✓ Criação de modelos capazes de se autocorrigir frente à possíveis problemas (SCHOLL *et al.*, 2012);

✓ Desenvolver métodos capazes de permitir a criação e replicação de modelos de maneira mais eficiente e automatizada (EYRE, SCOTT e FREEMAN, 2018).

Fonte: elaborado pelo próprio autor

4.2. Proposta de monitoramento dos modelos de simulação utilizados como Gêmeos Digitais

De forma a se estruturar a ferramenta de monitoramento dos modelos de simulação utilizados como GDs, sugere-se uma abordagem baseada em três fases principais: (i) Definição das variáveis de avaliação do GD (em que as variáveis-chave do sistema são selecionadas e preparadas) (ii) Criação e configuração da interface de monitoramento (a qual envolve a criação/configuração da interface que permite a classificação de dados físicos e do GD por meio do K-NN, e também a plotagem da acurácia do classificador na carta de controle p); e, por fim, (iii) Monitoramento periódico (onde a interface é atualizada periodicamente, ou seja, trata-se do funcionamento prático da ferramenta). Vale ressaltar também que, conforme descrito anteriormente, todo o desenvolvimento da proposta é pautado nas etapas do método de experimentação apresentadas no capítulo 3. Dessa forma, as seções a seguir apresentam as particularidades de cada uma das fases.

4.2.1. Definição das variáveis de avaliação

As variáveis de avaliação são os parâmetros do GD cruciais para o seu funcionamento, as quais possuem correspondentes tanto no modelo virtual quanto nos sistemas físicos. Sua escolha representa uma etapa crítica, uma vez que o sistema de monitoramento será construído e configurado a partir desta etapa. Nesse caso, é importante destacar que a escolha correta das variáveis de avaliação é tão importante quanto o funcionamento do sistema de monitoramento proposto. Segundo Sargent (2013), as variáveis de avaliação de um modelo de simulação estão relacionadas à finalidade do modelo, ou seja, aquelas associadas à tomada de decisão. Além disso, ao considerar o uso da simulação como GD, tal consideração é igualmente válida e, portanto, deve-se escolher as variáveis que impactam diretamente na tomada de decisão.

Considerando a estrutura de funcionamento do GD, entende-se que as variáveis de avaliação devem estar relacionadas à correspondência entre o ambiente físico e o virtual, ou seja, ligadas diretamente à integração entre os ambientes. Ao considerar GD de processos produtivos, algumas das variáveis de avaliação mais frequentes são tempos de processo, tempos de espera, taxa de utilização de recursos, produtividade, dentre outras.

Além disso, de acordo com a abordagem proposta, a escolha das variáveis deve ser em pares, ou seja, é importante garantir que cada variável esteja disponível tanto nos sistemas virtuais quanto nos físicos e na mesma proporção. Assim, é possível comparar os dois ambientes e inferir a validade do GD por meio da interface de monitoramento proposta. Nesse caso, juntamente com a definição das variáveis, é importante garantir a coleta periódica das mesmas, ou seja, deve-se sistematizar a coleta das variáveis no modelo virtual e também no ambiente físico.

Com relação ao ambiente virtual, é importante observar que as variáveis de avaliação estão relacionadas à saída do modelo e não aos dados de entrada, os quais também são coletados ao longo do tempo. Já com relação ao ambiente físico, as variáveis podem ser coletadas por meio de sensores, dispositivos IoT, bancos de dados, sistemas de gestão, dentre outros. Um grande diferencial da abordagem proposta é a possibilidade de se trabalhar com diversas variáveis de avaliação de forma simultânea, ou seja, diferentes variáveis advindas de ambos os sistemas podem ser analisadas e comparadas de uma só vez. Além disso, é possível também que tais variáveis possuam diferentes características, isto é, variáveis numéricas e categóricas, inteiras e reais e que se aderem ou não a distribuições padronizadas de probabilidade.

4.2.2. Criação e configuração da interface de monitoramento

Uma vez definidas as variáveis de avaliação, tal como descrito na seção anterior, o objetivo desta etapa é descrever o funcionamento da ferramenta de monitoramento proposta. Em primeiro lugar, a ferramenta foi planejada para ser uma interface amigável ao tomador de decisões, de forma a facilitar sua operação no dia a dia dos processos produtivos. Neste caso, optou-se por codificá-la em Python®, utilizando a biblioteca *Scikit-learn* (PEDREGOSA *et al.*, 2011). Além da classificação por meio do K-NN e da plotagem da sua acurácia na carta de controle, há inúmeras atividades auxiliares necessárias para o correto funcionamento do monitoramento do GD. Neste caso, as principais atividades consideradas são ilustradas pela Figura 4.8. Em primeiro lugar há a leitura e padronização dos dados, tanto físicos quanto digitais, os quais referem-se às variáveis de avaliação. Em seguida, há a classificação desses

dados por meio do K-NN seguida do cálculo da sua acurácia. Tal acurácia abastece um banco de dados visando a definição dos parâmetros da carta de controle, chamada de fase I, e, uma vez definidos, a ferramenta atua de maneira cíclica, apenas atualizando os pontos da carta a cada observação.

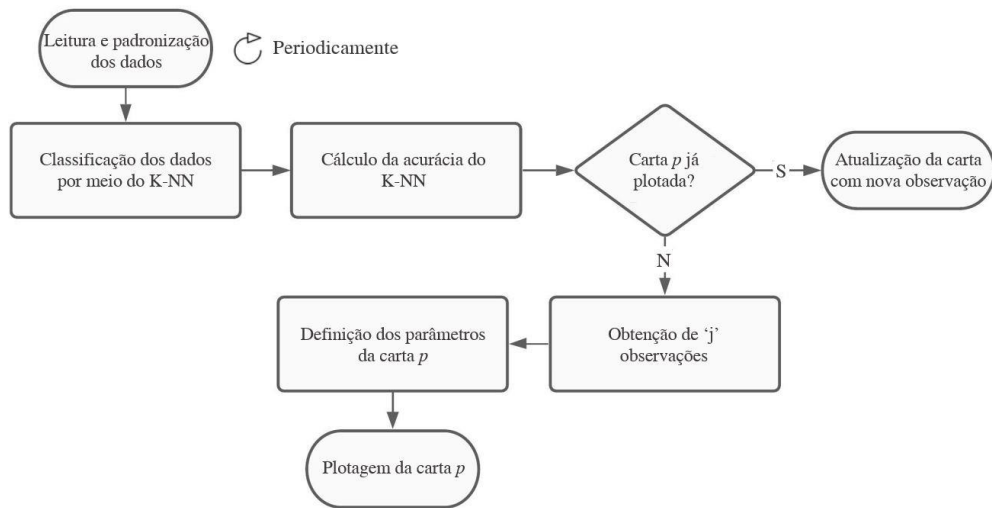


Figura 4.8 – Atividades desempenhadas pela ferramenta de monitoramento

Fonte: elaborado pelo próprio autor

Inicialmente, os dados do ambiente físico e do GD são coletados como um conjunto de dados, ambos nas mesmas proporções. Além disso, antes de serem classificados, é importante garantir que os dados estejam padronizados, uma vez que o sistema pode ter variáveis com diferentes características, ou seja, dados numéricos e categóricos, discretos e contínuos, reais e inteiros, entre outras combinações. Neste caso, várias técnicas de padronização podem ser adotadas e, para a ferramenta proposta, optou-se pelo chamado “*Robust Scaler*” (RAJU *et al.*, 2020). Primeiro, se houver variáveis categóricas, o algoritmo as transforma em numéricas. Em seguida, o *Robust Scaler* transformará todas as variáveis de avaliação, uma a uma, removendo a mediana de cada conjunto de dados e dimensionando os dados de acordo com o intervalo interquartil. Dessa forma, garantimos que todos os dados estejam na mesma escala. Além disso, é importante destacar que embora a aleatoriedade seja esperada em estudos de simulação, a presença de possíveis outliers nos dados não compromete a usabilidade da ferramenta proposta, uma vez que o *Robust Scaler* dimensiona os dados de acordo com a mediana do conjunto de dados (mais robusto para outliers do que a média).

Dessa forma, têm-se os dados prontos para serem classificados por meio do K-NN. Na etapa de classificação, o K-NN é primeiramente treinado a partir do conjunto de dados coletados anteriormente e por meio da Validação Cruzada. Nesse caso, adotou-se o parâmetro ‘Z’ = 5 (opção padrão da biblioteca utilizada), conforme sugerido por Li *et al.* (2019). Além disso, alguns testes com outros valores de ‘Z’ múltiplos de 5 (ou seja, 10, 15, 20, etc.) não resultaram em classificações mais precisas e, dessa forma, considerou-se ‘Z’=5 razoável para este estudo. Além disso, o valor ideal do parâmetro ‘K’ do classificador também foi definido a partir da Validação Cruzada, como sugerido por Zhang *et al.* (2021). Neste caso, para cada leitura do conjunto de dados (cada observação), o algoritmo testará alguns valores K e escolherá aquele com a maior acurácia de classificação. Hassanat, Abbad e Alhasanat (2014) afirmam que diversos trabalhos adotam $K = \sqrt{m}$, onde “m” é o tamanho do conjunto de dados, e, dessa forma, o algoritmo foi preparado para testar múltiplos deste valor.

Uma vez treinado, o K-NN está pronto para tentar separar os dados físicos e do GD. Nesse caso, se os dados advindos do modelo virtual são perfeitamente realistas, o K-NN não será capaz de distingui-los dos dados do sistema físico e, então, classificará cada observação de forma aleatória, resultando em uma acurácia de cerca de 50%. No entanto, se os dados do GD forem significativamente diferentes dos advindos do sistema físico, o classificador poderá classificá-los facilmente, resultando em uma acurácia de cerca de 100%. Em outras palavras, espera-se que a acurácia do K-NN aumente à medida que o modelo de simulação utilizado como GD se torna inválido. Portanto, uma acurácia do K-NN de 50% representa um GD válido (com 100% de correspondência com os sistemas físicos), enquanto uma acurácia de 100% do K-NN pode ser interpretada como um GD inválido (com 0% de correspondência com os sistemas físicos). Dessa forma, a tabela 4.1 apresenta mais detalhadamente a associação entre as acurácias do K-NN e a validade do GD, tal como considerada nessa abordagem.

Tabela 4.1 – Relação entre a acurácia do K-NN e a validade do Gêmeo Digital

Acurácia do K-NN	Correspondência do Gêmeo Digital com os sistemas físicos
50% - 55%	Muito forte (> 90%)
56% - 65%	Forte (70% - 90%)
66% - 75%	Moderada (50% - 70%)
76% - 85%	Fraca (30% - 50%)
85% - 100%	Muito fraca (0% - 30%)

Fonte: elaborado pelo próprio autor

É importante destacar que, embora o valor mínimo de acurácia do K-NN para considerar o modelo GD válido seja de 50%, conforme proposto anteriormente, espera-se que haja variabilidade durante a medição dessa acurácia, uma vez que se considerou o comportamento estocástico e aleatório dessa variável de monitoramento. Portanto, a carta de controle pode apresentar valores acima e abaixo de 50%, mas que devem estar dentro da faixa de controle definida para que se avalie a validade o modelo.

A acurácia do K-NN é então coletada ao longo do tempo e as primeiras 'j' observações (com tamanho amostral de 'n') servirão como base para a definição dos parâmetros de controle da carta p . Nesse caso, adotamos 'j' = 25, conforme sugerido por Montgomery (2009). Dessa forma, uma vez definidos seus parâmetros, torna-se possível plotar a carta p e usá-la para monitorar o GD ao longo do tempo. Essas duas etapas correspondem às fases I e II do estudo do gráfico de controle, respectivamente. Além disso, é importante mencionar que as atividades descritas são cíclicas e devem ser realizadas periodicamente e, portanto, após a plotagem da carta, esta será atualizada a cada nova observação. Além disso, conforme destacado anteriormente, a acurácia do K-NN desejada é em torno de 50%, o que indica que o GD está válido e apresenta resultados próximos o suficiente dos sistemas físicos, fornecendo diretrizes para a tomada de decisões com segurança. Neste caso, a carta p deverá permitir o monitoramento da acurácia do classificador ao longo do tempo e, caso algum elemento que compõe o GD sofra alguma falha funcional, espera-se que este efeito seja identificado pelo gráfico de controle.

Por fim, é importante destacar que o parâmetro 'm' (tamanho do conjunto de dados para a classificação K-NN) e 'n' (tamanho amostral considerado para a carta de controle p) são iguais. Neste caso, tal valor foi definido por meio da equação 4 apresentada na seção 2.5.2 e sugerida por Montgomery (2009). Para tal, considerou-se uma fração (p) da característica avaliada de 0,5 (valor alvo de acurácia do K-NN), a distância dos limites de controle da linha central (L) de 3 desvios padrão e, por fim, a magnitude do deslocamento do processo (δ) de 0,15 (visando garantir uma correspondência forte do GD com relação aos sistemas físicos, considerando a relação proposta na tabela 4.1). Portanto, obteve-se um tamanho amostral de 100 dados e, uma vez que o conjunto de dados utilizado pelo classificador é composto por dados físicos e virtuais (sempre em pares), a ferramenta de monitoramento foi configurada para coletar periodicamente 50 amostras de ambos os sistemas (físico e digital), totalizando uma amostra de tamanho 100.

4.2.3. Monitoramento periódico

O monitoramento periódico é o principal objetivo da abordagem proposta. Diferente da validação de projetos de simulação tradicionais, o objetivo não é garantir a validação do modelo durante sua fase de construção, mas sim em seu funcionamento ao longo do tempo no apoio à tomada de decisões. Além disso, outro fator a se levar em conta é que, enquanto a etapa de validação do modelo durante sua construção é realizada pelo próprio modelador e equipe envolvida, o monitoramento do modelo durante sua operação é de responsabilidade do tomador de decisões e, dessa forma, o foco principal da abordagem desenvolvida é a acreditação do modelo de simulação utilizado como GD. Neste caso, é importante definir algumas questões-chave quanto à utilização da ferramenta proposta, tal como frequência de monitoramento e ações considerando possíveis problemas quanto ao funcionamento do GD.

Por um lado, é importante ressaltar que a definição da frequência de monitoramento não depende necessariamente das características relacionadas ao nível de sincronismo do GD. Neste caso, conforme destacam Santos *et al.* (2022a), quando se trata do uso de simulação como GD, diversas abordagens podem ser consideradas, contemplando atualizações de modelos em tempo real ou quase real e intervalos de atualização constantes ou variáveis. No entanto, dada a facilidade de coleta de dados ao se considerar a operação dos GD e as soluções tecnológicas trazidas pela Indústria 4.0, sugere-se que a frequência do monitoramento esteja alinhada à capacidade de coleta dos dados, ou seja, na frequência em que o tamanho amostral é alcançado. Em outras palavras, os dados de ambos os sistemas, físico e digital, são coletados continuamente e cada vez que o tamanho da amostra de monitoramento ' n ' é atingido, todo o processo de monitoramento descrito na Figura 4.8 é realizado novamente, formando um processo cíclico.

Por outro lado, ao se considerar ações frente a possíveis problemas relacionados ao GD, ou seja, possíveis pontos fora dos limites de controle da carta p , é importante ressaltar que a ferramenta proposta foca exclusivamente no monitoramento e não no ajuste autônomo do sistema visando corrigi-lo. Isso porque, em casos onde o GD é considerado não válido, há a possibilidade de haver problemas que vão além do funcionamento do modelo, tais como falhas nos próprios sistemas físicos ou na estrutura de comunicação entre o modelo e o ambiente físico. Além disso, ao observar sinais fora de controle na carta p , indicando que o GD não está válido, é importante verificar se é um caso de alarme falso ou se há um problema a ser resolvido. Se for positivo para problemas em seu funcionamento, o usuário deve retornar à fase de construção do GD visando identificar a causa do problema, corrigi-lo e validar novamente o modelo antes de

utilizá-lo para apoio a decisões. Uma vez operando novamente, vale ressaltar que o GD deve ser monitorado novamente ao longo do tempo, reiniciando todo o ciclo proposto.

Nota-se, portanto, três estágios do GD quanto à sua validade: (I) Funcionando corretamente: onde a GD está operando conforme o esperado no suporte à tomada de decisões; (II) Verificação Necessária: onde a ferramenta de monitoramento proposta indica um ponto fora de controle e o tomador de decisão deve verificar possíveis falsos alarmes; e (III) Reajuste necessário: onde um ou mais problemas foram confirmados e o usuário precisa retornar à fase de construção da GD a fim de identificar os problemas e corrigi-los antes de se basear no GD para a tomada de decisões. A Figura 4.9 ilustra o ciclo de uso da ferramenta de monitoramento proposta, bem como os estágios do GD, conforme mencionado anteriormente.

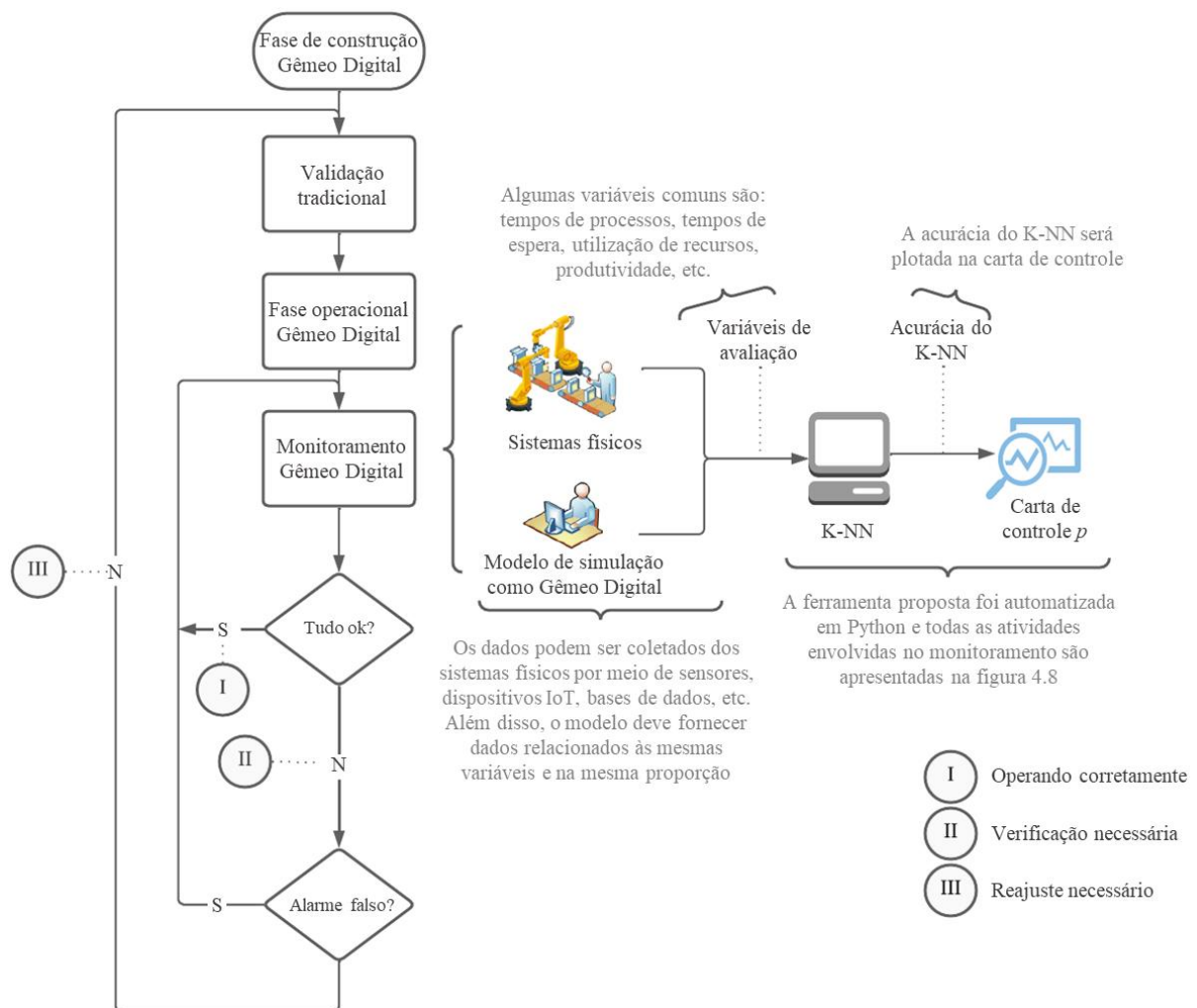


Figura 4.9 – Ciclo de uso da ferramenta de monitoramento e estágios do GD

Fonte: elaborado pelo próprio autor

4.3. Resultados Experimentais

A presente seção apresenta a aplicação da abordagem proposta de forma a se avaliar sua utilização no que tange o monitoramento do modelo de simulação como GD no decorrer do tempo. Neste caso, optou-se, inicialmente, pela adoção de objetos de estudo teóricos de forma a se verificar a aplicabilidade da ferramenta em situações onde o GD é sujeito a causas especiais capazes de impactar seus resultados. Entende-se que tal análise seria impraticável considerando objetos de estudo reais, uma vez que um GD já implementado em um caso real possui papel fundamental para apoio à tomada de decisões e, neste contexto, induzir causas especiais não seria viável. Em seguida, após demonstrar a funcionalidade da ferramenta de monitoramento, a mesma foi implementada em objetos de estudo reais, ou seja, em GDs já validados e prontos para apoiar a tomada de decisões. Neste caso, o objetivo é verificar a aderência da ferramenta proposta nos objetos de estudo, os quais possuem maior nível de detalhes e complexidade se comparados aos casos teóricos. Além disso, o intuito principal da abordagem proposta é a adição de um suplemento aos GDs pré-existentes de forma a torná-los ainda mais robustos e eficientes, capazes de não só auxiliar na tomada de decisões, como também se auto monitorar ao longo do tempo.

4.3.1. Objetos de estudo teóricos (Objetos I - V)

No caso dos objetos de estudo teóricos, optou-se por emular dados correspondentes aos ambientes físico e digital considerando distribuições de probabilidade padronizadas. Conforme destacado anteriormente, o objetivo neste caso é avaliar se a ferramenta proposta se comporta conforme o esperado quanto ao seu funcionamento, bem como explorar situações especiais onde os dados físicos diferem significativamente dos dados virtuais, fato que representaria um possível caso onde o GD não é válido. Optou-se pela adoção de cinco objetos de estudo teóricos com diferentes graus de complexidade, considerando as distribuições de probabilidade Lognormal, Bimodal, Poisson e Normal Multivariada com 3 e 5 variáveis, representando os objetos de estudo de I a V, respectivamente. Tal escolha se justifica por se tratarem de distribuições comumente utilizadas em projetos de simulação. Além disso, é importante ressaltar que o objetivo não é testar a proposta em todas as distribuições possíveis, mas sim demonstrar sua aplicabilidade. Neste caso, para cada objeto de estudo, adotou-se as três etapas referentes à abordagem proposta, ou seja, em primeiro lugar houve a seleção das variáveis de avaliação, seguida da criação e configuração da interface de monitoramento e, por fim, o monitoramento periódico. A Tabela 4.2 apresenta os parâmetros das distribuições adotadas.

Tabela 4.2 – Parâmetros das distribuições adotadas nos objetos de estudo teóricos

Objeto de estudo	Distribuição	Parâmetros
I	Lognormal	<i>Local = 0 e Escala = 1</i>
II	Binomial	50% <i>Normal</i> (100; 3) e 50% <i>Normal</i> (110; 10)
III	Poisson	<i>Lambda = 1</i>
IV	Normal	<i>Médias = [60; 500; 40] e</i>
	Multivariada (3 variáveis)	<i>Matriz de covariância = [[9; 60; 20], [60; 1600; 350], [20; 350; 100]]</i>
V	Normal	<i>Médias = [229,7; 249,0; 967,5; 964,2; 269,6] e</i>
	Multivariada (5 variáveis)	<i>Matriz de covariância = [[96,6; 63,7; 66,4; 67,2; 40,8], [63,7; 67,6; 52,8; 46,7; 42,7], [66,4; 52,8; 77,3; 53,4; 34,9], [67,2; 46,7; 53,4; 53,3; 35,4], [40,8; 42,7; 34,9; 35,4; 34,2]]</i>

Fonte: elaborado pelo próprio autor

Para cada estudo de caso foram emuladas 100 observações, cada uma com um tamanho de amostra $n = 100$ (conforme definido anteriormente). Tais observações representam o que seriam as variáveis de avaliação advindas do sistema físico e do modelo virtual, nas mesmas proporções, coletadas ao longo do tempo e armazenadas em diferentes conjuntos de dados já prontos para abastecer a interface de monitoramento. Nota-se que os objetos de estudo adotados apresentam diferentes características quanto às variáveis de avaliação. Neste caso, considerando os objetos de estudo I, II e III, cada conjunto de dados (físico e virtual) possui apenas uma coluna de dados que representa uma variável de avaliação (' x '), enquanto o objeto de estudo IV e V apresentam três (' x_1 ', ' x_2 ' e ' x_3 ') e cinco (' x_1 ', ' x_2 ', ' x_3 ', ' x_4 ' e ' x_5 ') colunas de dados, respectivamente. As Figuras 4.10 e 4.11 ilustram a comparação do que seriam as variáveis de avaliação advindas do ambiente físico e virtual, considerando os objetos de estudo I e IV, respectivamente. Vale destacar que o comportamento dos casos II e III é semelhante ao caso I, enquanto o caso V é semelhante ao caso IV, porém com mais duas variáveis de avaliação.

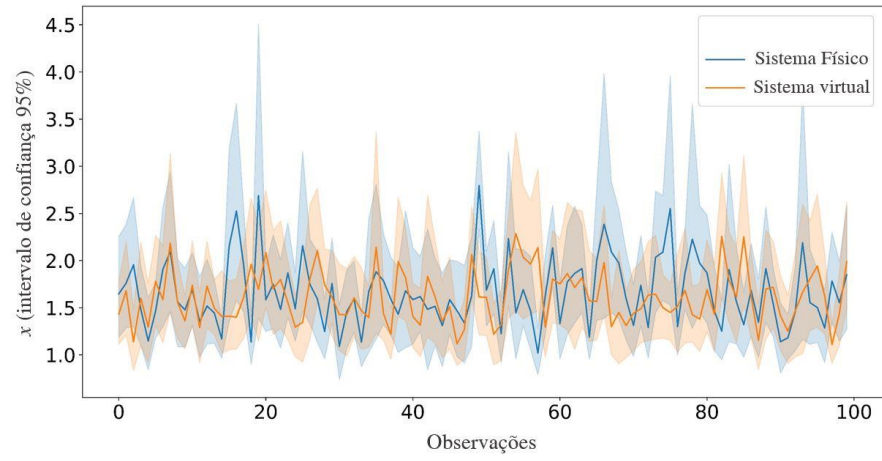


Figura 4.10 – Comparação dos dados físicos e virtuais (Objeto de estudo I)

Fonte: elaborado pelo próprio autor

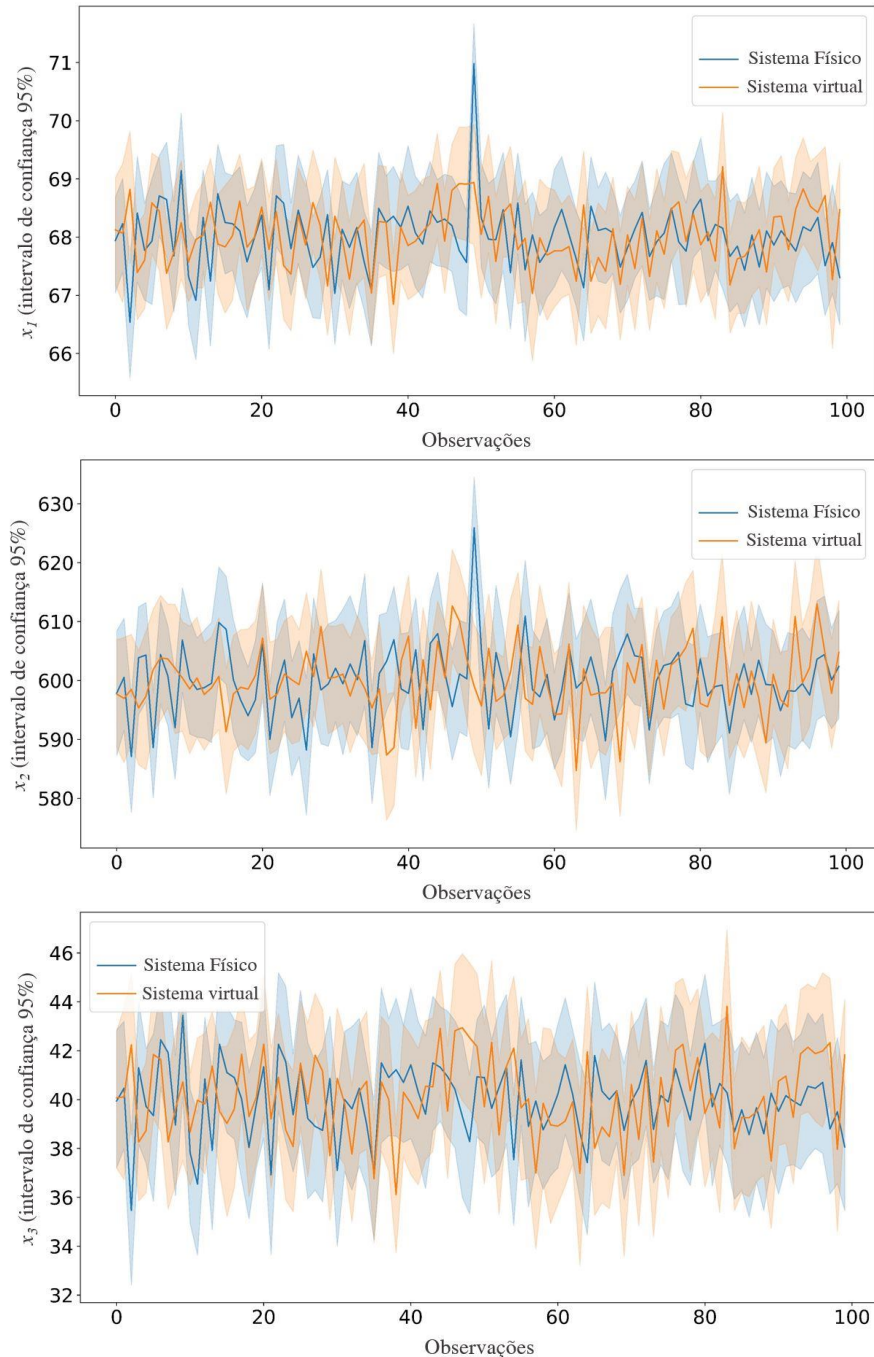


Figura 4.11 – Comparação dos dados físicos e virtuais (Objeto de estudo IV)

Fonte: elaborado pelo próprio autor

Conforme destacado anteriormente, a interface de monitoramento foi configurada para padronizar os dados e avaliar o valor do parâmetro 'K' do classificador por meio de Validação Cruzada, repetindo o procedimento a cada observação. A Figura 4.12 ilustra o procedimento para definição do valor de 'K' para uma determinada observação, o qual baseia-se na avaliação de $K = \sqrt{n}$ e seus múltiplos (onde n é o tamanho da amostra adotada) e considerando como valor ótimo aquele que resulta na maior acurácia do classificador. A cada observação tal

procedimento é repetido e obtém-se uma nova acurácia do K-NN. Além disso, conforme definido na abordagem proposta, as primeiras 25 observações foram utilizadas para construir os limites da carta de controle, enquanto as demais foram plotadas no gráfico, representando os dados coletados ao longo do tempo (fase I e II da carta, respectivamente).

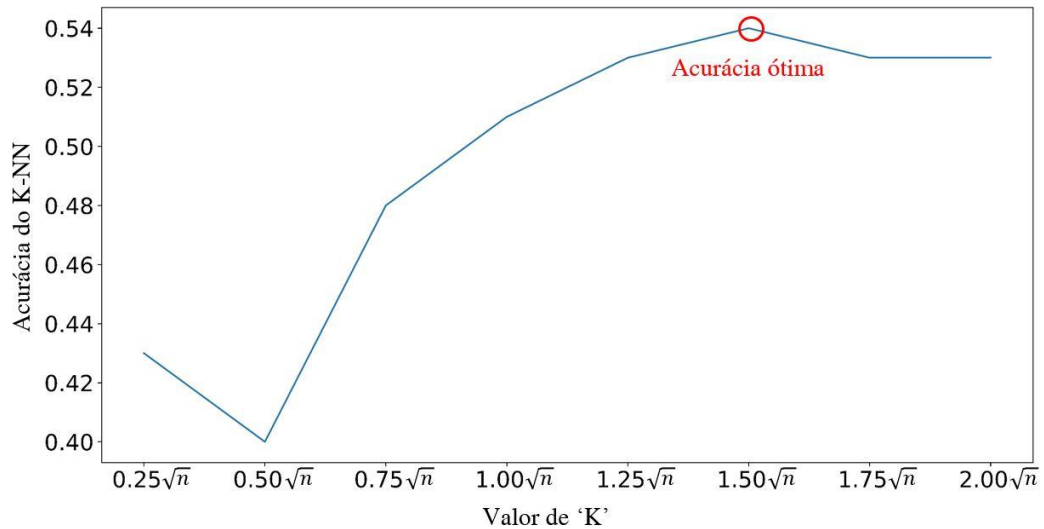


Figura 4.12 – Ilustração do procedimento adotado para obtenção do valor ótimo de 'K'

Fonte: elaborado pelo próprio autor

Outro ponto importante a ser destacado é com relação ao objetivo de se avaliar a abordagem proposta quando submetida a situações de possíveis problemas do GD. Neste caso, foi induzida uma causa especial em uma das observações (variação de 10% na observação nº 50 apenas no conjunto de dados que representaria os dados físicos) visando avaliar o desempenho da ferramenta de monitoramento na identificação de tal causa. É importante notar que uma diferença de 10% entre o ambiente físico e o virtual é comumente considerada comum em projetos de simulação. Entretanto, considerando que o uso da simulação como GD está associado a decisões de grande impacto, é importante monitorar seu funcionamento e identificar os sinais que podem estar causando tal variação. Dessa forma, as Figuras 4.13 a 4.17 apresentam as cartas de controle referente aos objetos de estudo I a V, respectivamente.

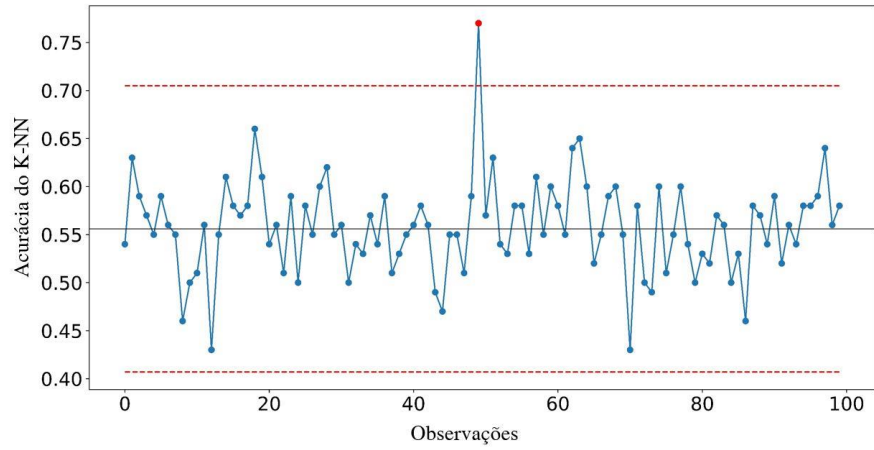


Figura 4.13 – Carta de controle para o objeto de estudo I

Fonte: elaborado pelo próprio autor

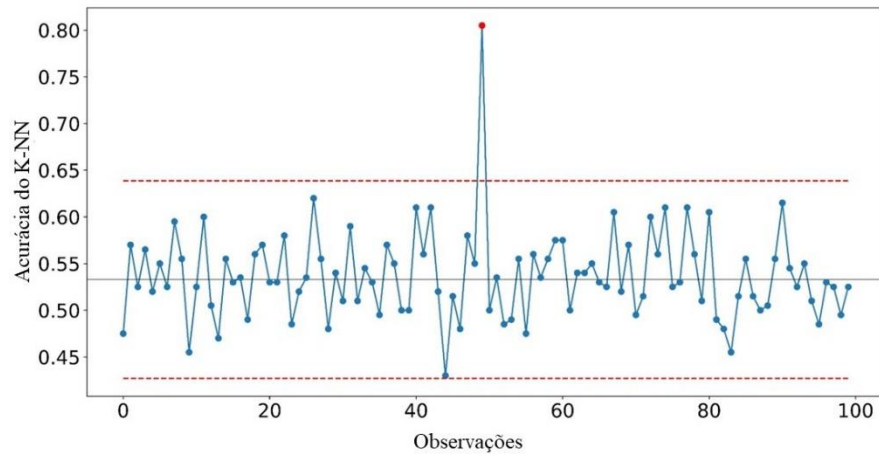


Figura 4.14 – Carta de controle para o objeto de estudo II

Fonte: elaborado pelo próprio autor

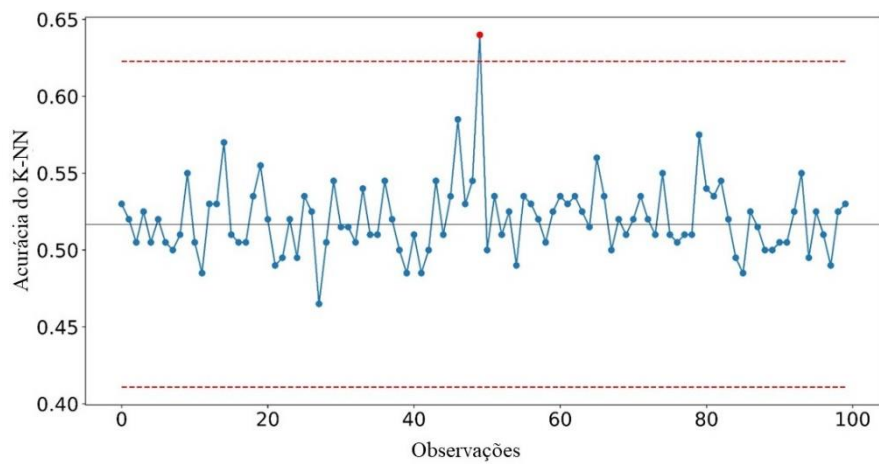


Figura 4.15 – Carta de controle para o objeto de estudo III

Fonte: elaborado pelo próprio autor

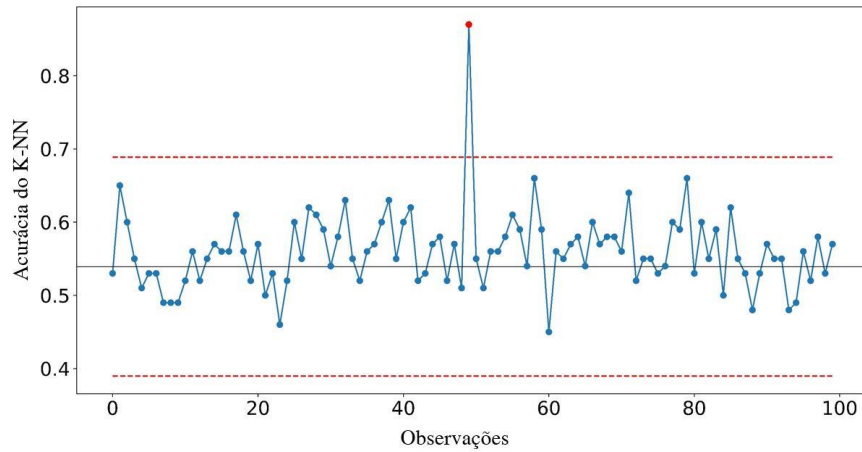


Figura 4.16 – Carta de controle para o objeto de estudo IV

Fonte: elaborado pelo próprio autor

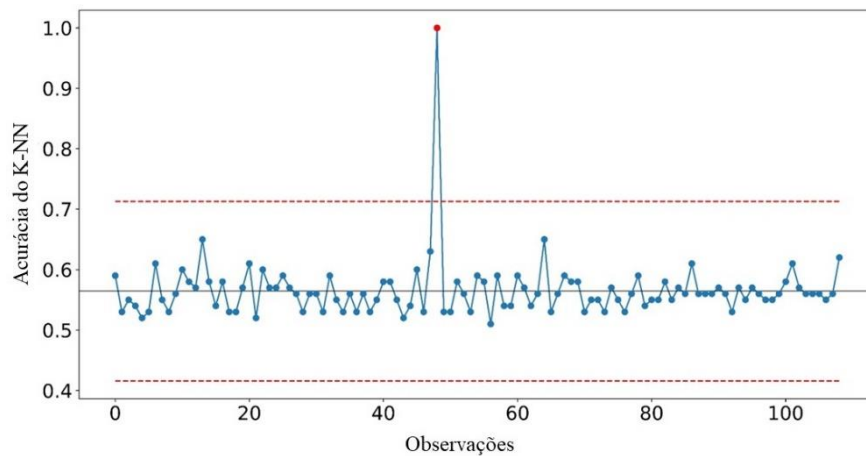


Figura 4.17 – Carta de controle para o objeto de estudo V

Fonte: elaborado pelo próprio autor

É possível notar que, para todos os objetos de estudo teóricos, a interface de monitoramento operou conforme o esperado. Em primeiro lugar, a acurácia média do K-NN para todos os casos representa uma correspondência muito forte ou forte do GD com os sistemas físicos, demonstrando que o classificador é capaz de identificar a semelhança entre o que seriam os dados virtuais e físicos, fato que era esperado já que se tratam de dados emulados. Além disso, ao considerar uma pequena variação no conjunto de dados que representaria as variáveis de avaliação advindas do ambiente físico, nota-se que o classificador conseguiu identificar essa causa especial, conforme demonstrado pelo sinal fora de controle nas cartas. Por fim, replicou-se as cartas de controle 1000 vezes para analisar seu comportamento, mantendo a causa especial na amostra nº 50. Neste caso, conforme mostrado na Figura 4.18, nota-se que em cerca de 97% das cartas plotadas foi possível identificar a causa especial (pontos fora de controle) na amostra

nº 50 considerando o objeto de estudo I. O mesmo foi observado para os demais objetos de estudo. Portanto, conclui-se, nesta primeira análise experimental, que a ferramenta proposta opera conforme o planejado e pode ser adotada em objetos de estudo reais.

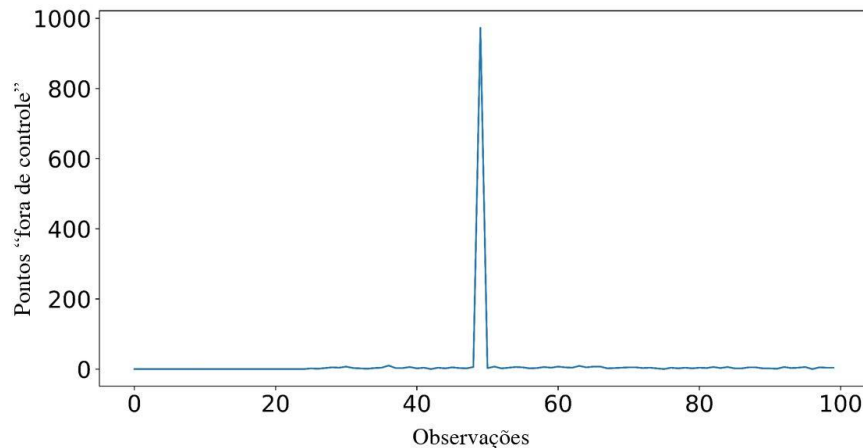


Figura 4.18 – Replicação da carta para identificar a causa especial (objeto de estudo I)

Fonte: elaborado pelo próprio autor

4.3.2. Objetos de estudo reais (Objetos VI e VII)

Considerando a aplicação da abordagem proposta em objetos de estudo reais, optou-se por dois GDs. A seguir são apresentados os casos e respectivos detalhes.

4.3.2.1. Objeto de estudo VI

O primeiro objeto de estudo refere-se a um GD que foi projetado para uma empresa de confecções *Fast Fashion* de médio porte. Tal segmento possui alta variabilidade de demanda, carecendo de ferramentas de decisão mais rápidas e eficientes. Neste caso, o GD foi planejado por meio de um modelo de SED para otimizar o planejamento operacional de uma das linhas de produção. A referida linha produz três tipos de produtos distintos (itens de vestuário), os quais têm sua demanda variando ao longo do ano, e o fluxo de produção é dividido em 8 postos de trabalho (Processo B ao Processo I), além das áreas de Recepção (A) e Expedição (J). O processo é majoritariamente manual e parte do transporte interno de produtos é realizado automaticamente por meio de um veículo guiado por automação (AGV). Vale ressaltar que, com exceção do processo B, o qual se trata de uma tarefa específica, a linha opera com operadores multifuncionais, ou seja, os demais processos são operados por operadores treinados para executar outras tarefas dentro do fluxo produtivo. Neste caso, os operadores podem ser

divididos em dois grupos, sendo o grupo I composto por até cinco operadores, responsáveis pelos processos C, D, E e G, e grupo II composto por até três operadores, responsáveis pelos demais processos (F, H, I e J). A Figura 4.19 ilustra a planta baixa da linha com o fluxo de produção e a figura 4.20 apresenta o modelo de SED em 3D construído por meio do *software* FlexSim®.

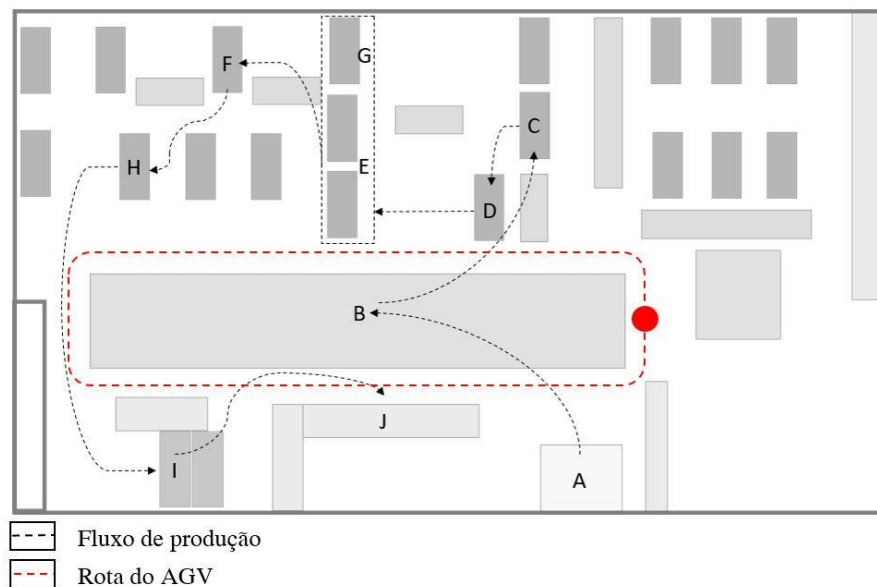


Figura 4.19 – Planta baixa e fluxo da linha utilizada como objeto de estudo VI

Fonte: elaborado pelo próprio autor



Figura 4.20 – Modelo de SED do objeto de estudo VI

Fonte: elaborado pelo próprio autor

O GD escolhido foi previamente construído, verificado e validado por meio das etapas apresentadas no capítulo 3 deste trabalho. Ou seja, o GD já se encontra operacional e, neste

caso, o objetivo é implementar a ferramenta de monitoramento proposta visando aprimorar o sistema de tomada de decisões. Trata-se de um GD que opera em tempo quase real e de forma não autônoma, ou seja, o modelo virtual é atualizado semanalmente e fornece orientações para a tomada de decisões. Além do modelo de SED, o GD baseia-se também em uma ferramenta de IA para previsão e uma *dashboard* de tomada de decisão. Nesse caso, o algoritmo IA, baseado em Redes Neurais Artificiais (RNA), prevê o comportamento da demanda semanal considerando o histórico de demanda disponível no banco de dados local. Além disso, com base na demanda prevista, o modelo de SED testa diferentes estratégias de planejamento de recursos semanais e indica a melhor decisão quanto ao dimensionamento de recursos (físicos e humanos). A *dashboard* de decisão integra os ambientes físico e virtual e fornece uma interface amigável para o tomador de decisão por meio de botões automatizados onde o usuário pode executar os modelos de IA e SED, além de visualizar as diretrizes para as decisões. A Figura 4.21 ilustra a estrutura de funcionamento do GD. Vale destacar que este objeto de estudo foi utilizado como base para outros trabalhos envolvendo o uso de GDs baseados em simulação e mais informações podem ser consultadas em (SANTOS *et al.*, 2021).

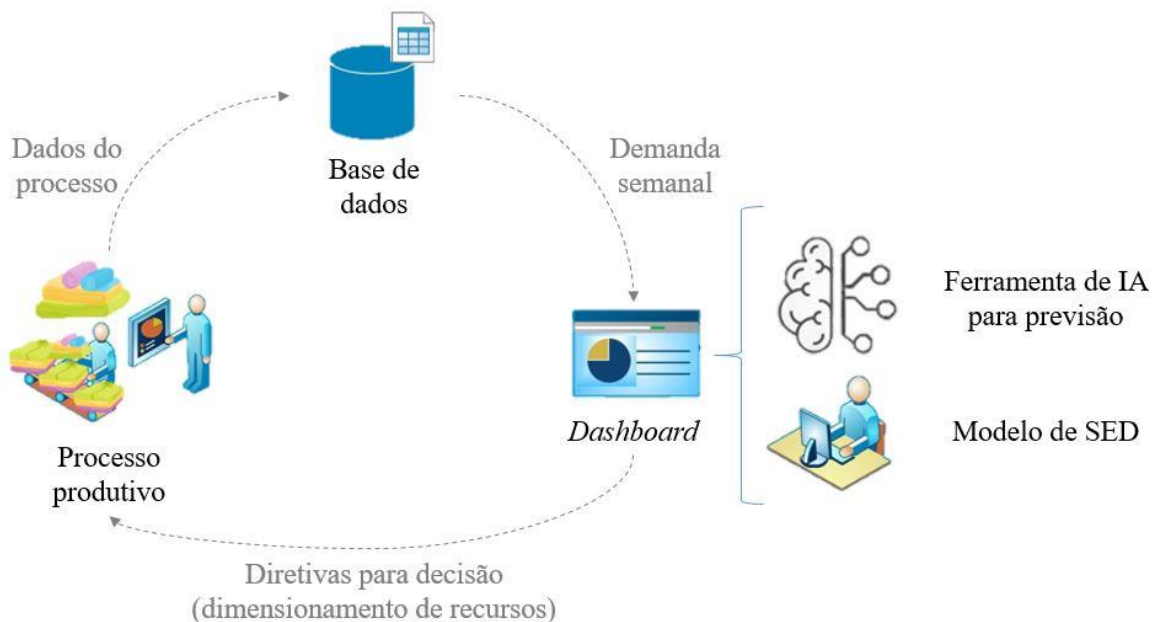


Figura 4.21 – Estrutura de funcionamento do Gêmeo Digital do objeto de estudo VI

Fonte: elaborado pelo próprio autor

Uma vez definido o objeto de estudo VI, o próximo passo é a implementação da abordagem proposta. Para tal, foram seguidas as três fases propostas na abordagem, ou seja, em primeiro lugar, selecionou-se as variáveis de avaliação do GD, seguida da configuração da ferramenta

de monitoramento e, por fim, foi permitida a avaliação periódica do GD. Para a seleção das variáveis de avaliação, optou-se pelos tempos de espera e tempo de processamento de cada produto, considerando os principais postos de trabalho (processo C a I) e também o tempo de transporte por AGV (totalizando 14 variáveis). Neste caso, além do histórico de demanda, o banco de dados local também possui dados de produção coletados por meio de etiquetas de radiofrequência, as quais captam informações sobre os produtos durante seu processamento. Dessa forma, o acesso aos dados advindos do sistema físico é garantido e, além disso, o modelo de SED também foi configurado previamente para oferecer, periodicamente, as mesmas informações referentes às variáveis de avaliação. Na medida em que são coletados, os dados (físicos e virtuais) são armazenados periodicamente em conjuntos de dados na própria base de dados local, estando prontos para a entrada na ferramenta de monitoramento.

A demanda média semanal da linha produtiva gira em torno de 750 itens de vestuário (incluindo os três tipos de produtos) e a ferramenta de monitoramento foi planejada para coletar os dados físicos em amostras de tamanho 50 durante a semana, enquanto os dados do modelo virtual também são coletados na mesma proporção, totalizando conjuntos de dados de tamanho 100 (tamanho da amostra definido inicialmente na seção 4.2.2). Assim, se a linha produzir 750 itens na semana, por exemplo, serão coletadas cerca de 15 amostras. Da mesma forma que nos casos teóricos, a ferramenta de monitoramento foi configurada para ser atualizada a cada amostragem, onde os dados são padronizados e o classificador K-NN é treinado. Nesse caso, a cada amostragem, foi obtida a acurácia do K-NN e, após o período correspondente a duas primeiras semanas de coleta de dados, foi possível definir os limites de controle a partir das primeiras 25 observações (fase I da carta de controle). Em seguida, as próximas amostras foram plotadas na carta de controle à medida que eram coletadas, permitindo o monitoramento periódico do GD. Para este trabalho, foi registrado o primeiro mês de operação da ferramenta de monitoramento, conforme ilustrado pela carta de controle apresentada na Figura 4.22.

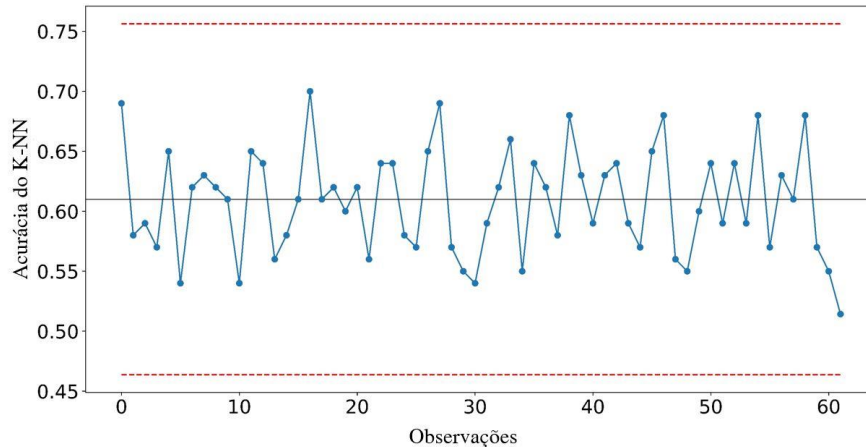


Figura 4.22 – Carta de controle para o objeto de estudo VI

Fonte: elaborado pelo próprio autor

Nota-se que, da mesma forma que nos casos teóricos, a interface de monitoramento operou conforme o esperado. Neste caso, a acurácia média do K-NN indica uma forte correspondência do GD com relação aos sistemas físicos. Outro ponto a se destacar é com relação à complexidade do GD adotado. Tratam-se de 14 variáveis de avaliação que não necessariamente seguem o comportamento de distribuições de probabilidade padronizadas e, mesmo assim, nota-se que a ferramenta de monitoramento é capaz de comparar os ambientes físico e virtual e identificar sua similaridade. A carta não indicou qualquer ponto fora do controle, uma vez que se tratou de um experimento controlado e sem nenhuma indução de causa especial.

4.3.2.2. Objeto de estudo VII

O segundo objeto de estudo refere-se a um GD implementado em uma célula de produção automatizada. Ao considerar uma linha automatizada, é importante garantir o correto funcionamento de todos os equipamentos e sistemas que a compõem. Neste contexto, o GD selecionado é baseado em um modelo SED e tem como foco o monitoramento de uma célula de produção. Esta célula possui uma estrutura de produção simples, composta por dois postos de trabalho que fabricam dois tipos de produtos, considerando três turnos de trabalho. Primeiramente, as matérias-primas chegam na área de chegada (A) e são enviadas para as áreas de abastecimento (B e D) que fornecem a matéria-prima para os postos de trabalho (C e E), respectivamente. O transporte entre as áreas é realizado por uma esteira automática e o deslocamento entre essas áreas e os postos de trabalho é realizado por braços robóticos (R1 e R2). As estações de trabalho são automatizadas do tipo de comando numérico computadorizado

(CNC). As Figuras 4.23 e 4.24 ilustram o fluxo de produção da célula e seu modelo de simulação 3D modelado em FlexSim®, respectivamente.

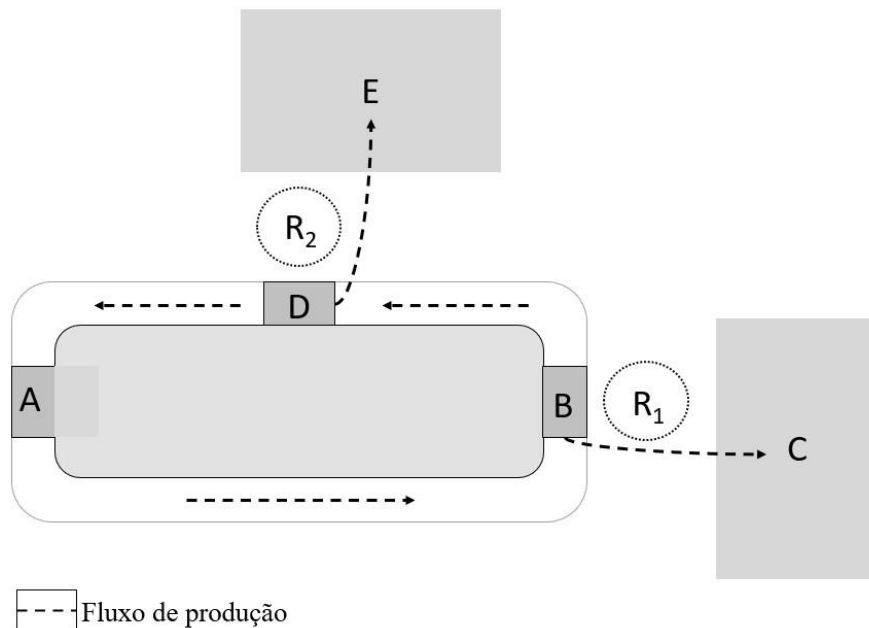


Figura 4.23 – Planta baixa e fluxo da linha utilizada como objeto de estudo VII

Fonte: elaborado pelo próprio autor

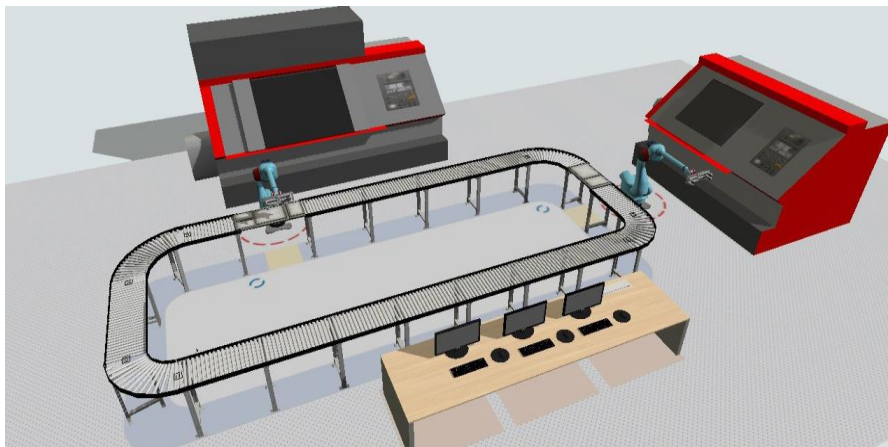


Figura 4.24 – Modelo de SED do objeto de estudo VII

Fonte: elaborado pelo próprio autor

Segundo Santos *et al.* (2022b), o uso de modelos virtuais para avaliar a performance de sistemas produtivos e analisar métricas durante sua operação é uma das principais aplicações dos GDs baseados em simulação. Neste caso, o GD escolhido foi planejado e implementado para dar suporte à célula de produção por meio de indicadores-chave de desempenho (KPIs), como *lead time* total, tempos de parada, taxa de produção, eficácia geral do equipamento, entre outros. Os

dados do processo são coletados em tempo real por meio de diversos sensores e dispositivos inteligentes e são armazenados na base de dados local. Além disso, a quantidade de peças em processamento (WIP) é utilizada para atualizar o modelo SED em um intervalo próximo de tempo real, com atraso de alguns minutos, espelhando a célula em um ambiente virtual e permitindo a comparação entre o comportamento esperado da célula e o real. Neste caso, têm-se uma abordagem não autônoma, uma vez que o GD é utilizado para avaliar o processo e fornecer uma base técnica para a tomada de decisões, porém o modelo não interfere diretamente nos sistemas físicos. Além disso, uma *dashboard* de decisão também é adotada para permitir a integração entre o modelo SED e os sistemas físicos que compõem a célula produtiva. Cabe ressaltar que o GD em questão foi previamente construído, verificado e validado e a Figura 4.25 ilustra sua arquitetura de operação.

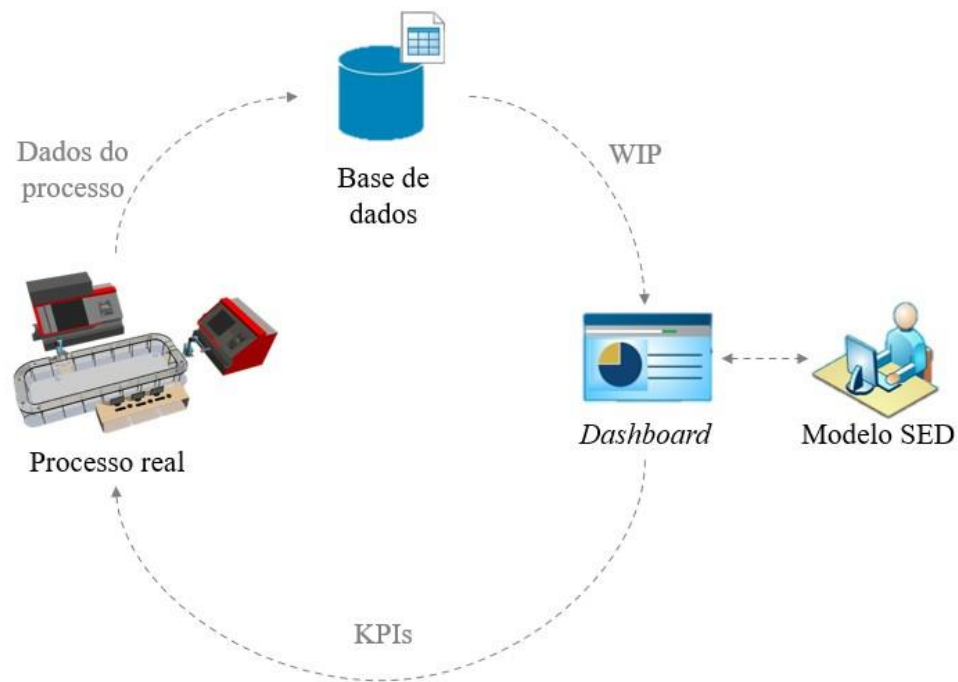


Figura 4.25 – Estrutura de funcionamento do Gêmeo Digital do objeto de estudo VII

Fonte: elaborado pelo próprio autor

De maneira análoga ao objeto de estudo VI, foram seguidas as três etapas propostas visando o monitoramento do GD. Com relação à seleção das variáveis de avaliação, optou-se também pelos tempos de espera e tempos de processamento de cada produto, considerando ambas as estações de trabalho C e E (totalizando 4 variáveis). Nesse caso, tais variáveis são coletadas continuamente por meio de sensores presentes na célula de produção, compondo os dados advindos dos sistemas físicos. Além disso, o modelo de simulação também oferece,

periodicamente, as mesmas informações do GD, ou seja, as mesmas variáveis de avaliação são coletadas do ambiente virtual na mesma proporção em que são coletadas do ambiente físico. Todas as variáveis de avaliação são armazenadas, na medida em que são coletadas, em diferentes conjuntos de dados, ficando disponíveis para o processo de monitoramento.

A demanda média de um turno de trabalho da célula produtiva é de cerca de 110 itens (incluindo os dois tipos de produtos fabricados) e a ferramenta de monitoramento foi planejada para coletar os dados físicos em amostras de 50 durante a jornada de trabalho, enquanto os dados virtuais também são coletados na mesma proporção, totalizando conjuntos de dados de tamanho 100, conforme definido anteriormente. Dessa forma, se a célula produzir 110 itens em um turno, cerca de duas amostras serão coletadas e inseridas à ferramenta de monitoramento. A cada amostragem, todo o processo descrito na seção 4.2.2 é realizado, onde inicialmente os dados são padronizados e o K-NN é treinado visando comparar os ambientes físico e virtual. Para o desenvolvimento deste trabalho, registrou-se observações de cerca de 50 turnos de trabalho, sendo as primeiras 25 observações utilizadas para definir os limites da carta de controle (fase I da carta) enquanto as demais foram utilizadas no processo de monitoramento periódico, ou seja, na atualização da carta de controle (fase II) e monitoramento do GD. A Figura 4.26 ilustra a carta de controle considerando este período de monitoramento.

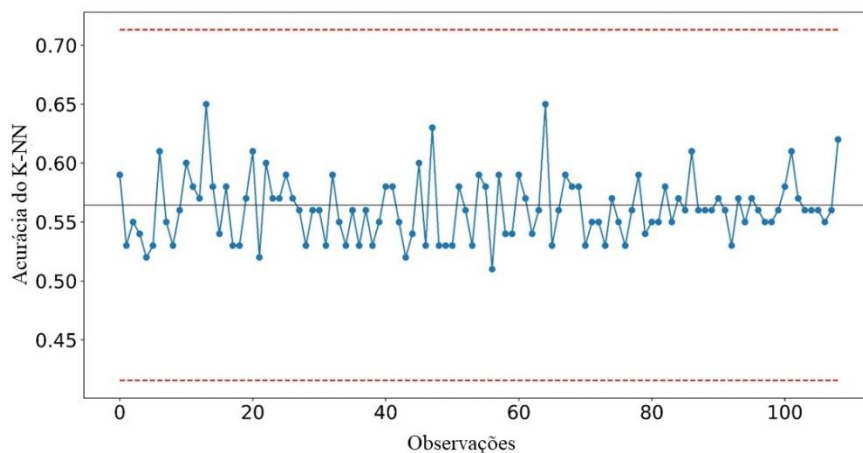


Figura 4.26 – Carta de controle para o objeto de estudo VII

Fonte: elaborado pelo próprio autor

Da mesma forma que nos demais casos apresentados, a ferramenta de monitoramento funcionou conforme o esperado. Nota-se que foi possível comparar o comportamento do GD com os sistemas físicos que compõem a célula de produção, identificando sua similaridade ao longo do tempo. Foi possível avaliar as quatro variáveis de avaliação de uma só vez e, neste caso, a

acurácia média do K-NN indica uma correspondência forte entre o modelo e o ambiente físico. Além disso, de maneira análoga ao objeto de estudo VI, não foi induzida qualquer causa especial no processo e, dessa forma, a carta de controle não indicou nenhum ponto fora de controle que pudesse indicar o mal funcionamento do GD e seus componentes principais.

4.3.2.3. Discussões finais acerca dos objetos de estudo reais (VI e VII)

Além disso, é importante avaliar o desempenho da carta de controle e, neste caso, adotou-se dois indicadores: (i) Qualidade da plotagem e (ii) *Average Run Length* (ARL). Conforme nota-se a partir das Figuras 4.22 e 44.26 (cartas de controle dos objetos de estudo VI e VII, respectivamente), as cartas conseguem representar de maneira clara e objetiva o comportamento do sistema, fato que atesta sua qualidade de plotagem. Além disso, com relação ao ARL ou, em outras palavras, à taxa de alarmes falsos, pôde-se avaliar os objetos de estudo separadamente utilizando a equação 5 apresentada na seção 2.5.2. Em primeiro lugar, o objeto de estudo VI obteve um ARL de cerca de 714, considerando um pr de 0,0014 (obtido a partir da distribuição Binomial e considerando os parâmetros do gráfico: $LC = 0,61$, $LSC = 0,76$ e $LIC = 0,46$, os quais foram obtidos com base nas equações 2 e 3). Tal resultado significa que, considerando que o processo está sob controle, um ponto fora de controle aparecerá a cada 714 observações. Da mesma forma, com relação ao objeto de estudo VII, obteve-se um ARL de cerca de 625, considerando um pr de 0,0016 (obtido da distribuição Binomial e considerando o gráfico parâmetros: $CL = 0,57$, $UCL = 0,72$ e $LCL = 0,42$, também obtidos com base nas equações 2 e 3). Em outras palavras, se o processo estiver em controle, um ponto fora de controle aparecerá a cada 625 observações. Segundo Aebtarm e Bouguila (2011), um ARL de 370 costuma ser referência na literatura considerando três níveis de desvio padrão para os limites do gráfico e, portanto, a abordagem proposta mostra-se viável. Além disso, o usuário pode ainda analisar o impacto no desempenho da carta (por meio do ARL) nos casos em que seus parâmetros são alterados. Vale destacar que o ARL possui uma escala contínua e que pode variar ao longo da utilização da carta, ou seja, os valores apresentados referem-se a observações pontuais e que serviram para avaliar o desempenho da carta na abordagem proposta.

Por fim, é importante destacar que a abordagem proposta traz simplicidade ao monitoramento. No objeto de estudo VI, todas as 14 variáveis de avaliação foram analisadas de uma só vez, enquanto o objeto de estudo VII foi avaliado considerando suas 4 variáveis também de uma só vez. Logo, a ferramenta proposta simplifica o processo de monitoramento e avaliação dos GDs com apenas um parâmetro (acurácia K -NN), o qual é plotado na carta de controle

periodicamente, independentemente da complexidade dos modelos utilizados para compor o GD. Neste contexto, trata-se de um grande diferencial, visto que em métodos tradicionais de validação, cada variável de avaliação deve ser comparada aos pares, resultando em um monitoramento mais desafiador e que pode comprometer o uso do GD na tomada de decisões caso o modelo seja atualizado em intervalos de tempo curtos. Por exemplo, os objetos de estudo VI e VII exigiriam cerca de 14 e 4 testes de hipóteses simultâneos e periódicos, respectivamente. Além disso, pode haver a necessidade de adotar diferentes testes dependendo da natureza das variáveis de avaliação (ou seja, testes paramétricos e não paramétricos). Por fim, considerando múltiplos testes de hipóteses, o nível de confiança pode ser prejudicado, pois há um efeito combinado do erro associado a cada teste. Em contraste, a abordagem de monitoramento proposta integra confiabilidade e agilidade no monitoramento de modelos de GD baseados em simulação.

Além disso, realizou-se uma análise comparativa para demonstrar as vantagens da abordagem proposta em relação às práticas atuais baseadas em testes periódicos de hipóteses. Neste caso, considerou-se o mesmo período de monitoramento descrito anteriormente: um mês de operação para o objeto de estudo VI e 50 turnos de trabalho para o estudo de caso VII. Primeiro, verificou-se a normalidade dos dados para escolher o teste de hipótese adequado e eles não se ajustavam à distribuição normal (*P-value* do teste de Anderson-Darling $< 0,05$). Em seguida, a cada amostra coletada (composta por dados dos ambientes físico e virtual), foi realizado o teste de Mann-Whitney para cada variável de avaliação para comparar os dados físicos e virtuais. Nesse caso, se o GD estiver funcionando conforme o esperado, espera-se que o teste apresente um *P-value* $\geq 0,05$. A Figura 4.27 apresenta a avaliação dos objetos de estudo VI e VII por meio dessa abordagem e optou-se por apresentar uma variável de avaliação de cada estado apenas para fins ilustrativos.

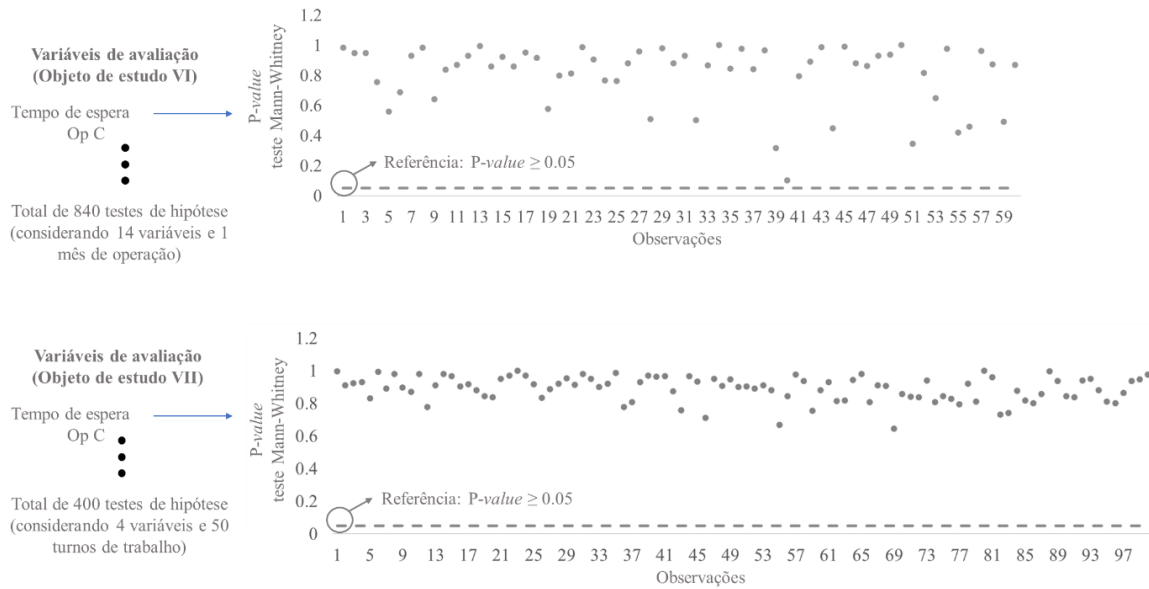


Figura 4.27 – Avaliação dos objetos de estudo VI e VII por meio de testes de hipóteses periódicos

Fonte: elaborado pelo próprio autor

Nota-se que, embora tenha sido possível avaliar as GDs ao longo do tempo por meio de testes de hipóteses periódicos, essa abordagem pode ser impraticável ou inviável. Para o objeto de estudo VI, foram necessários cerca de 840 testes para avaliar o GD durante um mês de operação e considerando todas as 14 variáveis de avaliação. Da mesma forma, para o objeto de estudo VII, foram necessários cerca de 400 testes de hipóteses para permitir a avaliação do GD durante 50 turnos de trabalho e considerando todas as 4 variáveis de avaliação. Conforme destacado anteriormente, essa abordagem pode se tornar ainda mais complicada se houver a necessidade de diferentes testes de hipóteses ou a inclusão de mais variáveis de avaliação. Por outro lado, a ferramenta proposta permite a avaliação dos GDs de forma mais simples e rápida, conforme demonstrado neste trabalho.

Além disso, se considerarmos características como correlação entre as variáveis, o monitoramento seria ainda mais complexo, exigindo tratamentos adicionais para o monitoramento (MONTGOMERY, 2009). Assim, se uma determinada variável apresentar um comportamento inesperado (resultado de uma causa especial), as demais variáveis correlacionadas serão impactadas, resultando em conclusões equivocadas sobre a validade do GD. Por outro lado, a abordagem proposta não se baseia em análises individuais, mas no monitoramento de todas as variáveis de forma integrada, tornando o monitoramento mais ágil. Além disso, a ferramenta de monitoramento pode apresentar a matriz de correlação das variáveis de avaliação, indicando ao tomador de decisão os níveis de correlação entre as variáveis e facilitando a investigação de possíveis ocorrências de causas especiais caso o

processo esteja fora de controle. Por exemplo, verifica-se no objeto de estudo VI que há variáveis de avaliação correlacionadas, conforme mostra a Figura 4.28. Neste caso, nota-se que as variáveis “Tempo de processo C” e “Tempo de espera D” possuem forte correlação.

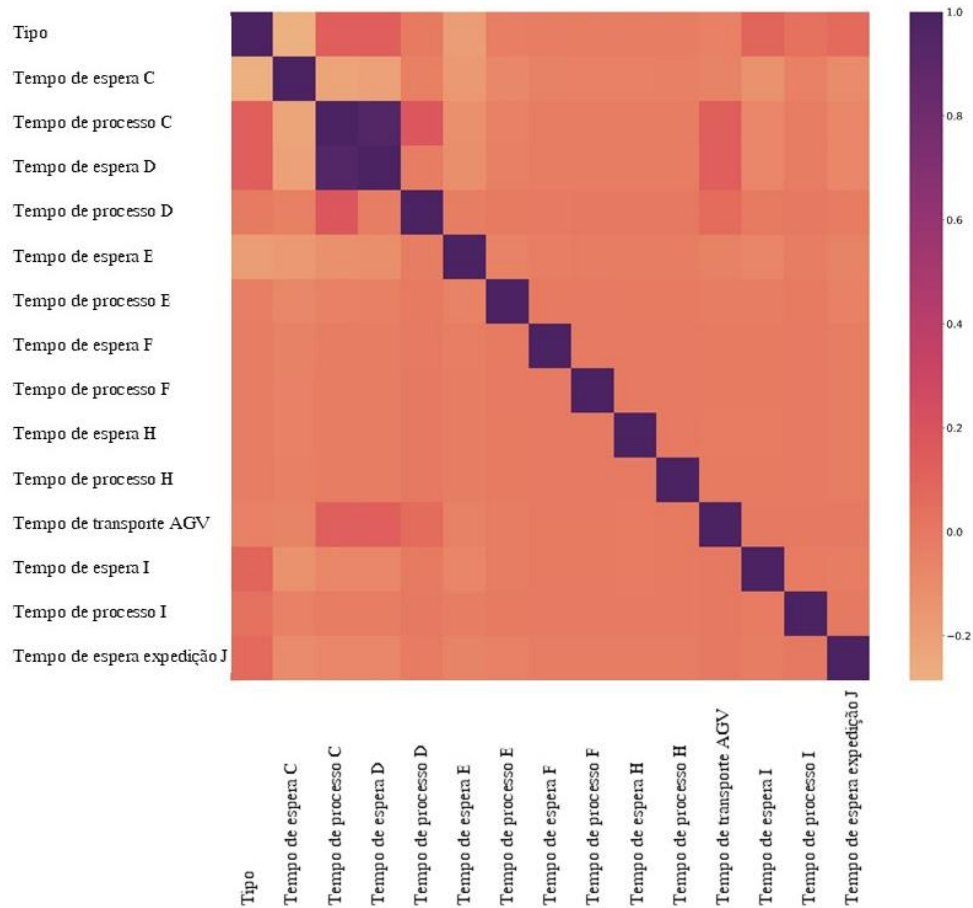


Figura 4.28 – Mapa de correlação entre as variáveis do objeto de estudo VI

Fonte: elaborado pelo próprio autor

5. CONCLUSÕES

Uma vez que a abordagem proposta foi apresentada e testada em objetos de estudo visando avaliar sua aplicabilidade e eficiência, pode-se destacar as conclusões deste trabalho, com ênfase nos resultados principais, discussões, limitações e oportunidades futuras. Sendo assim, as próximas seções irão discutir tais tópicos, apresentando primeiramente a síntese dos resultados, seguida das limitações da pesquisa e, por fim, recomendações para trabalhos futuros.

5.1. Síntese dos Resultados

O uso da simulação como GD representa uma alternativa flexível e eficiente para a tomada de decisão. A partir da integração do modelo de simulação com o ambiente físico (ou seja, máquinas, dispositivos inteligentes e bancos de dados), é possível espelhar os processos por meio de uma cópia virtual inteligente e altamente sincronizada, visando a otimização das decisões. No entanto, apesar da ampla aplicabilidade desta abordagem e dos inúmeros trabalhos publicados nesta área nos últimos anos, nota-se que ainda existem alguns desafios e oportunidades. Nesse contexto, destaca-se a necessidade de monitorar o funcionamento do modelo de simulação como GD ao longo do tempo, uma vez que este é atualizado periodicamente de acordo com as mudanças físicas. Além da validação tradicional durante a fase de construção do modelo, é necessário avaliar sua validade ao longo do tempo, considerando que possíveis problemas como falhas de comunicação, paradas não programadas, alterações nos processos e falhas do modelo podem prejudicar o correto funcionamento do GD, impactando a tomada de decisão.

Uma vez que as abordagens tradicionais de validação não consideram a atualização constante do modelo de simulação e considerando que há uma lacuna na literatura de trabalhos focados neste propósito, há necessidade de se criar técnicas e soluções voltadas para o monitoramento de modelos de simulação utilizados como GD. A partir de uma RSL, pôde-se realizar uma varredura da literatura acerca do uso da simulação como GD a fim de justificar o desenvolvimento deste trabalho. Dessa forma, esta tese propôs uma abordagem baseada em Aprendizado de Máquina e carta de controle para permitir a avaliação de modelos de simulação como GD durante sua fase operacional. Mais precisamente, utilizou-se do classificador K-NN para comparar o comportamento do modelo de simulação com os sistemas físicos e, por meio da carta de controle p , foi possível monitorar a validade do modelo GD ao longo do tempo.

A abordagem proposta é baseada em três etapas principais, começando com a seleção das variáveis de avaliação a serem monitoradas. Em seguida, há a etapa de construção /configuração da interface de monitoramento, a qual compara os dados dos sistemas físicos com os dados do ambiente virtual. Na sequência, por meio da acurácia do classificador K-NN, é possível inferir sobre a validade do GD, sendo esta plotada na carta p ao longo do tempo, possibilitando a avaliação temporal do modelo. Cabe salientar que essa última etapa corresponde ao monitoramento constante e periódico do GD por meio da ferramenta desenvolvida. Nesse caso, destaca-se a praticidade da ferramenta na avaliação de modelos, permitindo o monitoramento de maneira mais rápida e eficiente mesmo em casos mais complexos.

É importante destacar que a abordagem proposta traz simplicidade na avaliação da validade do modelo GD. Neste caso, independentemente da quantidade e do(s) tipo(s) de variáveis de avaliação, características relacionadas à complexidade do GD, a ferramenta simplifica o processo de monitoramento com apenas um parâmetro (acurácia K -NN), o qual é plotado na carta de controle periodicamente. Neste contexto, trata-se de um grande diferencial, visto que em métodos tradicionais de validação, cada variável de avaliação deve ser comparada aos pares, ou seja, cada variável do modelo comparada com sua correspondente do ambiente físico, resultando em um monitoramento mais desafiador e que pode comprometer o uso do GD na tomada de decisões. Além disso, pode haver a necessidade de se adotar diferentes testes dependendo da natureza das variáveis de avaliação e, ao se considerar múltiplos testes, o nível de confiança pode ser prejudicado já que há um efeito combinado do erro associado a cada teste. Em contraste, a abordagem de monitoramento proposta integra confiabilidade e agilidade no monitoramento de modelos de GD baseados em simulação.

Além disso, trata-se de uma abordagem flexível a qual traz maior robustez aos Gêmeos Digitais baseados em simulação. Por um lado, o uso de simulação como Gêmeos Digitais representa uma alternativa flexível e que independente de pacotes comerciais ou equipamentos que integram o ambiente produtivo. Além disso, a simulação permite a personalização dos resultados do GD de acordo com as necessidades do tomador de decisões. Por outro lado, há a condição de que as decisões oriundas do GD geralmente impactam significativamente nos sistemas produtivos e, dessa forma, a confiabilidade do modelo é um fator crítico. Tal fator embasa e justifica o presente trabalho, uma vez que a abordagem proposta opera como um suplemento aos GDs visando torna-los mais robustos e confiáveis ao longo de sua utilização no apoio à decisão.

A abordagem proposta foi inicialmente testada em cinco casos teóricos visando avaliar sua aplicabilidade. Nesse caso, dados foram emulados representando ambos os sistemas, virtual e físico, de acordo com distribuições de probabilidade padronizadas, sendo a Lognormal, Bimodal, Poisson e Normal Multivariada com 3 e 5 variáveis, representando os objetos de estudo de I a V, respectivamente. Dessa forma, foi possível observar a funcionalidade da ferramenta a partir de experimentos nos quais o sistema foi submetido a estresse a fim de avaliar sua capacidade de identificar possíveis problemas com relação ao funcionamento do Gêmeo Digital. Logo, a partir da inserção de causas especiais que representariam um possível problema no funcionamento do GD, foi possível avaliar a capacidade da ferramenta proposta em identificar tais perturbações. É importante destacar que o objetivo desta primeira rodada de

testes não era testar a ferramenta proposta em todas as distribuições possíveis, mas sim demonstrar sua aplicabilidade a partir de distribuições comuns em projetos de simulação.

Além disso, a abordagem proposta foi implementada em dois objetos de estudo reais. Neste caso, diferente dos casos teóricos, o objetivo não foi avaliar o comportamento da ferramenta de monitoramento quando sujeita a perturbações e efeitos propositais, mas sim compreender sua capacidade de atuar como um suplemento de GDs reais visando torna-los mais robustos e eficientes. O primeiro objeto de estudo (objeto de estudo VI) trata-se de um GD já operacional planejado para uma empresa de médio porte do ramo de confecções. Trata-se de um GD não autônomo e que apoia decisões semanais relacionadas ao planejamento operacional da linha de produção. Já o segundo estudo (objeto de estudo VII) refere-se a um GD também já implementado e que apoia decisões em uma célula de produção automatizada. Também se têm um GD que opera próximo de tempo real, com um atraso de minutos com relação ao sistema físico e que é focado na avaliação de KPIs da célula, ou seja, o GD é não autônomo e apenas fornece diretivas para a tomada de decisões.

A partir da implementação da ferramenta de monitoramento nos casos teóricos e reais, ressaltase que a ferramenta se mostrou capaz de monitorar o funcionamento do GD e identificar possíveis causas especiais que podem comprometer seus resultados. Nos casos teóricos, a ferramenta foi capaz de identificar causas especiais que pudessem estar associadas ao mau funcionamento do GD. Por outro lado, nos casos reais, pôde-se perceber a versatilidade e flexibilidade da abordagem proposta, sendo compatível com diferentes abordagens de GD, contemplando modelos com diferentes características de conexão, integração e complexidade. Portanto, conclui-se que o objetivo deste trabalho foi cumprido, uma vez que se preenche uma lacuna na literatura quanto à validação de modelos de simulação utilizados como GD, propondo uma ferramenta capaz de avaliar periodicamente e sistematicamente os modelos quanto à sua validade. Por fim, destaca-se também o papel deste trabalho na disseminação do uso de GDs baseados em simulação no suporte aos sistemas de decisões, uma vez que se propõe a adição de um suplemento flexível e versátil aos modelos já implementados visando trazer mais credibilidade e robustez à tomada de decisões.

5.2. Limitações da pesquisa

É importante destacar algumas limitações que fazem parte do contexto desta pesquisa. Neste caso, em primeiro lugar, o leitor deve considerar que a abordagem apresentada se refere a uma proposta de ferramenta que visa preencher uma lacuna teórica e prática acerca da validade dos

modelos GD baseados em simulação. Logo, não faz parte do escopo desta pesquisa apresentar uma abordagem com a melhor eficiência e eficácia possível, fato que demandaria o teste com diferentes técnicas de AM como classificadores, bem como diferentes cartas de controle para o monitoramento. A seção a seguir apresenta tais testes como uma proposta para trabalhos futuros. Além disso, o foco deste trabalho limita-se a GDs baseados em modelos de simulação, não englobando GDs com base em outras ferramentas e pacotes comerciais. Porém, vale ressaltar que tal fato não limita o uso da ferramenta proposta nestes outros tipos de GD e, neste caso, também foram sugeridos trabalhos nesta linha de pesquisa. Por fim, este trabalho foi desenvolvido considerando aplicações e testes envolvendo GDs de processos e não de produtos, fato que poderia impactar na utilização da ferramenta proposta. Tal abordagem também foi destacada como possível trabalho futuro.

5.3. Recomendações para trabalhos futuros

Pode-se destacar algumas oportunidades de pesquisa com foco em aprimorar a abordagem proposta e expandir seu escopo de atuação. A seguir são listadas tais oportunidades:

- Melhorias quanto ao funcionamento da ferramenta: a abordagem proposta, tal como descrita neste trabalho, permite ao tomador de decisões monitorar a validade do GD durante sua operação e identificar a ocorrência de possíveis problemas. Entretanto, entende-se como um ponto crítico quanto ao seu uso a identificação da causa raiz destes problemas. Neste caso, propõe-se, para pesquisas futuras, o aprimoramento da abordagem proposta de forma a permitir que o usuário, além de identificar falhas no funcionamento do GD que culminam na perda de validade, possa também identificar a causa raiz destas falhas. Dessa forma, ao retornar para a fase de construção/ validação do modelo, espera-se que o modelador seja capaz de rapidamente solucionar os eventuais problemas. Tal proposta vai ao encontro dos chamados “Gêmeos Digitais Auto Adaptativos”, os quais detectam falhas e auxiliam o usuário na correção das mesmas, um tema promissor quanto a pesquisas na área (SANTOS *et al.*, 2022b).
- Uma vez que a base da abordagem proposta é a comparação entre os ambientes físico e virtual por meio de um classificador K-NN e o monitoramento de sua acurácia a partir da carta de controle p , entende-se como uma oportunidade para trabalhos futuros o aprimoramento da ferramenta por meio de outros classificadores e cartas de controle. Neste caso, pode-se destacar diversas outras técnicas de AM capazes de classificar dados, tais como *Decision Tree* (ABDUL-RAHMAN *et al.*, 2021), *Artificial Neural*

Networks (AIZENBERG *et al.*, 2021), *Random Forest* (ALI, 2021), dentre outras. Além disso, a depender do formato das acurácias desses classificadores, pode-se utilizar diversas outras cartas de controle, tais como *X-barra & R*, *X- barra & S*, *I-MR*, *C*, *EWMA*, *CUSUM*, *cartas multivariadas*, dentre outras (COSTA, EPPRECHT e CARPINETTI, 2018; MONTGOMERY, 2009). Dessa forma, futuros trabalhos podem explorar possíveis melhorias quanto à capacidade de monitoramento da ferramenta proposta em busca de soluções cada vez mais eficientes.

- Além disso, entende-se como uma oportunidade para próximos trabalhos a aplicação da ferramenta proposta em outros contextos, com GDs focados em diferentes objetivos. Em especial, nota-se oportunidades relacionadas a GDs focados em apoiar decisões associadas ao planejamento de manutenção e setups de equipamentos, por exemplo, tal como destacado em Santos *et al.* (2023). Além disso, o presente trabalho focou em GDs de processos produtivos, logo, pode-se explorar o uso da abordagem proposta visando monitorar a validade de GDs de produtos. Neste caso, há características importantes que diferem os modelos de processos e de produtos, tais como as fases de modelagem, tipos de conexão, grau de sincronismo, dentre outras. Além disso, GDs baseados em pacotes comerciais ou outras ferramentas além da simulação também são candidatos para se aplicar a ferramenta proposta, visto que todos os testes realizados neste trabalho focam em GDs baseados em modelos de simulação.

APÊNDICE A - Publicações

i. Periódicos

Publicados/aceitos

SANTOS, C. H. *et al.* Digital Twin simulation models: a validation method based on machine learning and control charts. **International Journal of Production Research**, v. x, p. xx–xx, 2023.

SANTOS, C. H. *et al.* Simulation-based digital twins monitoring: an approach focused on models' accreditation. **The International Journal of Advanced Manufacturing Technology**, v. 124, p. 2423–2435, 2023.

CAMPOS, A. T. GABRIEL, G. T; SANTOS, C. H. *et al.* Integrating Computer Simulation and the Normalized Normal Constraint method to plan a temporary hospital for COVID-19 patients. **Journal of the Operational Research Society**, v. 74:2, p. 562–573, 2023.

PINTO, W. G. M.; MONTEVECHI, J. A. B.; MIRANDA, R. C.; SANTOS, C. H. *et al.* Optimisation via Simulation Applied to Reverse Logistics: a Systematic Literature Review. **International Journal of Simulation Modelling**, v. 22, p. 29–40, 2023.

OLIVEIRA, M. S.; SANTOS, C. H. *et al.* FaMoSim: a facilitated discrete event simulation framework to support online studies. **Production**, v. 33, p. 1-15, 2023.

SANTOS, C. H. *et al.* Decision support in productive processes through DES and ABS in the Digital Twin era: a systematic literature review. **International Journal of Production Research**, v. 60:8, p. 2662–2681, 2022.

SANTOS, C. H. *et al.* Use of Simulation in the Industry 4.0 context: creation of a Digital Twin to optimize decision making on non-automated process. **Journal of Simulation**, v. 16:3, p. 284–297, 2022.

AMARAL, J. V. S.; MIRANDA, R. C.; MONTEVECHI, J. A. B.; SANTOS, C. H. *et al.* Data envelopment analysis for algorithm efficiency assessment in metamodel-based simulation optimization. **The International Journal of Advanced Manufacturing Technology**, v. 121, p. 7493–7507, 2022.

AMARAL, J. V. S.; MIRANDA, R. C.; MONTEVECHI, J. A. B.; SANTOS, C. H. *et al.* Metamodeling-based simulation optimization in manufacturing problems: a comparative study. **The International Journal of Advanced Manufacturing Technology**, v. 120, p. 5205–5224, 2022.

CAMPOS, A. T.; SANTOS, C. H. *et al.* Safety assessment for temporary hospitals during the COVID-19 pandemic: a simulation approach. **Safety Science**, v. 147, p. 1–10, 2022.

SANTOS, C. H. *et al.* Decision-making in a fast fashion company in the Industry 4.0 era: a Digital Twin proposal to support operational planning. **The International Journal of Advanced Manufacturing Technology**, v. 116, p. 1653-1666, 2021.

SANTOS, C. H. *et al.* A decision support tool for operational planning: a Digital Twin using simulation and forecasting methods. **Production**, v. 30, p. 1–17, 2020.

SANTOS, C. H. *et al.* Application of the Discrete Event Simulation integrated with Lean concepts for implementing improvements in a military field hospital. **Revista Pesquisa Naval**, v. 31, p. 2–13, 2019.

Submetidos/em revisão

AMARAL, J. V. S.; SANTOS, C. H. *et al.* Digital Twin-based decisions in the energy sector: a systematic literature review. **Renewable and Sustainable Energy Reviews**, 2023.

PEREIRA, A. B. M. MONTEVECHI, J. A. B.; PINTO, W. G. M.; SANTOS, C. H. Simulation and Digital Twins to support Reverse Logistics decisions: a review. **International Journal of Simulation Modelling**, 2023.

AMARAL, J. V. S.; MIRANDA, R. C.; MONTEVECHI, J. A. B.; SANTOS, C. H. *et al.* Adaptive Metamodeling-based Simulation Optimization. **Computers and Industrial Engineering**, 2023.

LÚCIO, J. S.; MIRANDA, R. C.; SANTOS, C. H. Bringing together Lean Six Sigma and Discrete Event Simulation: a systematic literature review. **Journal of Simulation**, 2023.

ii. Livros

Capítulos de livros publicados/aceitos

SANTOS, C. H.; MONTEVECHI, J. A. B. **Digital Twins Architecture**. In: Lv, Z.; Fersman, E. (Org.). **Digital Twins: Basics and Applications**. Springer, 2022, p. 1.

SANTOS, C. H. **Utilização da simulação na Indústria 4.0: proposta de um sistema cyberfísico**. In: Ernane Rosa Martins. (Org.). **Engenharia de Produção: gestão de qualidade, produção e operações**. 1.ed. Editora Científica Digital, 2021, v. 2, p. 195.

iii. Congressos

Publicados/aceitos

MONTEVECHI, J. A. B. *et al.* **Using Generative Adversarial Networks to validate Discrete Event Simulation models**. Proceeding of 2022 Winter Simulation Conference. **Anais...**Singapura: 2022.

OLIVEIRA, M. S. *et al.* **Facilitated Discrete Event Simulation framework to support online studies in industry**. Simpósio Brasileiro de Pesquisa Operacional 2022. **Anais...**Juiz de Fora, Brasil: 2022.

SANTOS, C. H. *et al.* **Monitoramento de modelos de simulação usados como Gêmeos Digitais de processos: uma abordagem baseada em Machine Learning e Carta de Controle**. Encontro Nacional de Engenharia de Produção 2022. **Anais...**Foz do Iguaçu, Brasil: 2022.

GABRIEL, G. T. *et al.* **Discriminação de dados para validação de modelos computacionais utilizando Redes Neurais Artificiais**. Encontro Nacional de Engenharia de Produção 2022. **Anais...**Foz do Iguaçu, Brasil: 2022.

PINTO, W. G. M. *et al.* **Otimização via simulação aplicado a logística reversa: uma Revisão Sistemática de Literatura**. Encontro Nacional de Engenharia de Produção 2022. **Anais...**Foz do Iguaçu, Brasil: 2022.

AMARAL, J. V. S. *et al.* **Redes Neurais para Otimização via Simulação por Metamodelagem: uma aplicação no setor industrial**. Encontro Nacional de Engenharia de Produção 2022. **Anais...**Foz do Iguaçu, Brasil: 2022.

CAMPOS, A. T. *et al.* **Redes Adversárias Generativas: uma nova abordagem para modelagem de dados de entrada em projetos de simulação**. Encontro Nacional de Engenharia de Produção 2022. **Anais...**Foz do Iguaçu, Brasil: 2022.

MONTEVECHI, J. A. B. *et al.* **Input Data Modeling: An Approach using Generative Adversarial Networks**. Proceeding of 2021 Winter Simulation Conference. **Anais...**Online, USA: 2021.

SOUZA, J. M. *et al.* **Uso da Simulação integrada aos Conceitos Enxutos visando melhorar o desempenho de um processo no setor público federal**. Simpósio Brasileiro de Pesquisa Operacional 2021. **Anais...**Online, Brasil: 2021.

VENÂNCIO, A. M. *et al.* **Proposta de um Serious Game para apoio ao ensino de conceitos ligados ao Lean Office.** Encontro Nacional de Engenharia de Produção 2021. *Anais...Online*, Brasil: 2021.

AMARAL, J. V. S. *et al.* **Otimização Baseada em Metamodelos: Uma Abordagem para Metamodelagem em Simulação a Eventos Discretos.** Encontro Nacional de Engenharia de Produção 2021. *Anais...Online*, Brasil: 2021.

LIMA, Y. T. *et al.* **Aprendizagem de máquina para apoio no monitoramento de postos de trabalho.** Simpósio de Engenharia de Produção 2021. *Anais...Online*, Brasil: 2021.

MONTEVECHI, J. A. B. *et al.* **A Method Proposal for Conducting Simulation Projects in Industry 4.0: a Cyber-Physical System in an Aeronautical Industry.** Proceedings of 2020 Winter Simulation Conference. *Anais...Orlando, USA*: 2020.

GABRIEL, G. T. *et al.* **Utilização da Otimização, Simulação a Eventos Discretos e Design of Experiment para a tomada de decisão em um problema de agendamento de cirurgias.** Simpósio Brasileiro de Pesquisa Operacional 2020. *Anais...Online*, Brasil: 2020.

SANTOS, C. H. *et al.* **Utilização da Simulação integrada aos Conceitos Enxutos e princípios da Indústria 4.0: uma proposta de Gêmeo Digital para planejamento operacional em uma indústria têxtil.** Encontro Nacional de Engenharia de Produção 2020. *Anais...Online*, Brasil: 2020.

GABRIEL, G. T. *et al.* **Utilização da Simulação integrada à Otimização: uma proposta visando o planejamento operacional de um centro cirúrgico. In: XL Encontro Nacional de Engenharia de Produção.** Encontro Nacional de Engenharia de Produção 2020. *Anais...Online*, Brasil: 2020.

CAMPOS, A. T. *et al.* **Resposta à pandemia de Covid-19: planejamento e otimização de uma unidade de pronto atendimento utilizando simulação computacional.** Encontro Nacional de Engenharia de Produção 2020. *Anais...Online*, Brasil: 2020.

SANTOS, C. H. *et al.* **Proposta de um método para condução de projetos de simulação no contexto da Indústria 4.0: uma aplicação da simulação como sistema cyber-físico em uma indústria do setor aeronáutico.** Simpósio Brasileiro de Pesquisa Operacional 2019. *Anais...Limeira*, Brasil: 2019.

ALVIM, P. P. *et al.* **Utilização da simulação na avaliação dos resultados esperados pela implantação dos Conceitos Enxutos em um laboratório de análises clínicas.** Simpósio Brasileiro de Pesquisa Operacional 2019. *Anais...Limeira*, Brasil: 2019.

SIQUEIRA, L. B. *et al.* **Utilização da simulação como ferramenta de auxílio ao ensino de Conceitos Enxutos em ambientes hospitalares.** Simpósio Brasileiro de Pesquisa Operacional 2019. *Anais...Limeira*, Brasil: 2019.

FERREIRA, B. S. *et al.* **Aplicação da modelagem e simulação computacional como ferramenta para implementação da Produção Enxuta em uma indústria de fornos industriais.** Simpósio Brasileiro de Pesquisa Operacional 2019. *Anais...Limeira*, Brasil: 2019.

SOUZA, V. P. *et al.* **Aplicação dos Conceitos e Ferramentas da Manufatura Enxuta para Eliminação de Desperdícios e Redução de Custos em uma Empresa de Manufatura.** Encontro Nacional de Engenharia de Produção 2019. **Anais...**Santos, Brasil: 2019.

SANTOS, C. H. *et al.* **Utilização da simulação dentro do contexto da indústria 4.0: proposta de um sistema cyber-físico visando a redução de desperdícios em um processo de abastecimento de materiais de uma indústria aeronáutica.** Encontro Nacional de Engenharia de Produção 2019. **Anais...**Santos, Brasil: 2019.

SILVA, H. R. *et al.* **Aplicação dos conceitos e ferramentas Lean para redução do Lead Time de produção e entrega de uma empresa com tipologia de produção Make to Order.** Encontro Nacional de Engenharia de Produção 2019. **Anais...**Santos, Brasil: 2019.

HUGO, A. A. *et al.* **Proposta de modelagem dos processos de um almoxarifado através da técnica BPMN: uma aplicação em um hospital público do Sul de Minas.** Encontro Nacional de Engenharia de Produção 2019. **Anais...**Santos, Brasil: 2019.

VILAS BOAS, J. L. *et al.* **Desenvolvimento e Aplicação de um Roadmap para o Planejamento e Execução de Eventos Kaizen.** Encontro Nacional de Engenharia de Produção 2019. **Anais...**Santos, Brasil: 2019.

ROARK, G. *et al.* **Modelaje y Simulación computacional con FlexSim de un proceso de despacho y expedición en una industria cementera argentina.** Encontro Nacional de Engenharia de Produção 2019. **Anais...**Santos, Brasil: 2019.

SANTOS, C. H. *et al.* **Aplicação da Simulação a Eventos Discretos integrada com Conceitos Enxutos para a implementação de melhorias em um Hospital Militar de Campo.** Simpósio de Pesquisa Operacional e Logística da Marinha 2019. **Anais...**Rio de Janeiro, Brasil: 2019.

REFERÊNCIAS

ABBAS, Z.; NAZIR, H.; AKHTAR, N.; ABID, M.; RIAZ, M. On designing an efficient control chart to monitor fraction nonconforming. **Quality and Reliability Engineering**, v. 36, n. 2, p. 1–18, 2019.

ABDUL-RAHMAN, S.; ARIFIN, N.; HANAFIAH, M.; MUTALIB, S. Customer Segmentation and Profiling for Life Insurance using K-Modes Clustering and Decision Tree Classifier. **International Journal of Advanced Computer Science and Applications**, v. 12, n. 9, p. 434–444, 2021.

AEBTARM, S.; BOUGUILA, N. An empirical evaluation of attribute control charts for monitoring defects. **Expert Systems with Applications**, v. 38, n. 6, p. 7869–7880, 2011.

AIZENBERG, I.; BELARDI, R.; BINDI, M.; GRASSO, F.; MANETTI, S.; LUCHETA, A.; PICCIRILLI, M. A neural network classifier with multi-valued neurons for analog circuit fault diagnosis. **Electronics (Switzerland)**, v. 10, n. 3, p. 1–18, 2021.

- AKHAVIAN, R.; BEHZADAN, A. H. An integrated data collection and analysis framework for remote monitoring and planning of construction operations. **Advanced Engineering Informatics**, v. 26, n. 4, p. 749–761, 2012.
- AKHAVIAN, R.; BEHZADAN, A. H. Evaluation of queuing systems for knowledge-based simulation of construction processes. **Automation in Construction**, v. 47, p. 37–49, 2014.
- AKHAVIAN, R.; BEHZADAN, A. H. Coupling human activity recognition and wearable sensors for data-driven construction simulation. **Journal of Information Technology in Construction**, v. 23, p. 1–15, 2018.
- ALAM, K. M.; SADDIK, A. EL. C2PS: A Digital Twin Architecture Reference Model for the Cloud-Based Cyber-Physical Systems. **IEEE Access**, v. 5, p. 2050–2062, 2017.
- ALI, A. B. W. Prediction of Employee Turn Over Using Random Forest Classifier with Intensive Optimized Pca Algorithm. **Wireless Personal Communications**, v. 119, n. 4, p. 3365–3382, 2021.
- ALTAF, M. S.; LIU, H.; AL-HUSSEIN, M.; YU, H. **Online simulation modeling of prefabricated wall panel production using RFID system**. Proceedings of the 2015 Winter Simulation Conference. **Anais...**Huntington Beach, CA: 2015.
- ANAGNOSTOPOULOS, D.; NIKOLAIDOU, M. Timing Issues and Experiment Scheduling in Faster-Than-Real-Time Simulation. **Simulation**, v. 79, n. 11, p. 613–625, 2003.
- ANTONS, O.; ARLINGHAUS, J. C. Designing decision-making authorities for smart factories. **Procedia CIRP**, v. 93, p. 316–322, 2020.
- APSEMIDIS, A.; PSARAKIS, S.; MOGUERZA, J. M. A review of machine learning kernel methods in statistical process monitoring. **Computers & Industrial Engineering**, v. 142, p. 1–12, 2020.
- ASHRAFIAN, A.; PETTERSEN, O.; KUNTZE, K.; FRANKE, J.; ALFNES, E.; HENRIKSEN, K.; SPONE, J. **Full-Scale Discrete Event Simulation of an Automated Modular Conveyor System for Warehouse Logistics**. Proceedings of International Conference on Advances in Production Management Systems 2019. **Anais...**Austin, USA: 2019.
- AUGUSTO, V.; MURGIER, M.; VIALON, A. **A Modelling and Simulation framework for intelligent control of emergency units in the case of major crisis**. Proceedings of the 2018 Winter Simulation Conference. **Anais...**Gothenburg, Sweden: 2018.
- AVCI, O.; ABDELJABER, O.; KIRANYAZ, S.; HUSSEIN, M.; GABBOUJ, M.; INMAN, D. A review of vibration-based damage detection in civil structures: From traditional methods to Machine Learning and Deep Learning applications. **Mechanical Systems and Signal Processing**, v. 147, p. 1–45, 2021.
- BALCI, O. A life cycle for modeling and simulation. **Simulation**, v. 88, n. 7, p. 870–883, 2012.

BANKS, J. **Handbook of Simulation: Principles, Methodology, Advances, Applications and Practice**. 1. ed. Atlanta: Georgia Institute of Technology, 1998.

BANKS, J.; CARSON II; NELSON, B.; NICOL, D. **Discrete Event System Simulation**. 5. ed. New Jersey: Pearson Prentice Hall, 2010.

BARCIK, J. P.; MÖLLER, D. P. F.; VAKILZADIAN, H. **Cyber physical system application in transportation: Analysis and multiplatform implementation of a highway tollbooth case study**. Proceedings IEEE International Conference on Electro Information Technology 2016. **Anais...**Grand Forks, USA: 2016.

BARLAS, P.; HEAVEY, C. Automation of input data to discrete event simulation for manufacturing: A review. **International Journal of Modeling, Simulation, and Scientific Computing**, v. 07, n. 01, 2016.

BEREGI, R.; SZALLER, Á.; KÁDÁR, B. Synergy of multi-modelling of process control. **IFAC-PapersOnLine**, v. 51, n. 11, p. 1023–1028, 2018.

BERGMANN, S.; STELZER, S.; STRASSBURGER, S. **Initialization of simulation models using CMSD**. Proceedings of the 2011 Winter Simulation Conference. **Anais...**Phoenix, USA: 2011.

BILLER, B.; GUNES, C. **Introduction to simulation input modeling**. Proceedings of the 2010 Winter Simulation Conference. **Anais...** Baltimore, Maryland: 2010.

BOONTASRI, K.; TEMDEE, P. Determining Significant Classification Factors for Senior Learning: A Case Study of Thai Seniors and Social Media Skill Learning. **Wireless Personal Communications**, v. 115, p. 2951–2970, 2020.

BOSCHERT, S.; ROSEN, R. **Digital Twin - The Simulation Aspect**. In HEHENBERGER, P.; BRADLEY, D. (Eds.). **Mechatronic Futures: Challenges and Solutions for Mechatronic Systems and Their Designers**. [s.l.] Springer, 2016. p. 59–74.

BOTTANI, E.; MURINO, T.; VESPOLI, S. **From the Cyber-Physical System to the Digital Twin: the process development for behaviour modelling of a Cyber Guided Vehicle in M2M logic**. Proceedings of XXII Summer School Francesco Turco Industrial Systems Engineering 2017. **Anais...**2017.

BRADFORD, R.; SIMMONDS, R.; UNGER, B. **A Parallel Discrete Event IP Network Emulator**. Proceedings 8th International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems. **Anais...**California, USA: 2000.

BRAGLIA, M.; GABRIELLI, R.; FROSOLINI, M.; MARRAZZINI, L.; PADELLINI, L. **Using RFID technology and Discrete-Events, Agent-Based simulation tools to build Digital-Twins of large warehouses**. 2019 IEEE International Conference on RFID Technology and Applications (RFID-TA). **Anais...**Pisa, Italy: 2019.

CASSETTARI, L.; BENDATO, I; MOSCA, M.; MOSCA, R. Energy Resources Intelligent Management using on line real-time simulation: A decision support tool for sustainable manufacturing. **Applied Energy**, v. 190, p. 841–851, 2017.

CHEN, H.; HUROYAN, V.; SONI, U.; LU, Y.; MACIEJEWSKI, R.; KOBOUROV, S. Same Stats, Different Graphs: Exploring the Space of Graphs in Terms of Graph Properties. **IEEE Transactions on Visualization and Computer Graphics**, v. 27, n. 3, p. 2056–2072, 2019.

CHINNATHAI, M. K. ALKAN, B.; VERA, D.; HARRISON, R. Pilot to Full-Scale Production: A Battery Module Assembly Case Study. **Procedia CIRP**, v. 72, p. 796–801, 2018.

CHO, M.; SONG, M.; YOO, S.; REIJERS, H. An Evidence-Based Decision Support Framework for Clinician Medical Scheduling. **IEEE Access**, v. 7, p. 15239–15249, 2019.

CHOI, S.; KANG, G. Towards development of cyber-physical systems based on integration of heterogeneous technologies. **International Journal of Computer Applications in Technology**, v. 58, n. 2, p. 129–136, 2018.

CHONG, C. S.; SIVAKUMAR, A. I. **Design, development and application of an object oriented simulation toolkit for real-time semiconductor manufacturing scheduling**. Proceedings of the 2002 Winter Simulation Conference. **Anais...**San Diego, CA: 2002.

CHUKHROVA, N.; JOHANNSEN, A. Hypergeometric p -chart with dynamic probability control limits for monitoring processes with variable sample and population sizes. **Computers & Industrial Engineering**, v. 136, p. 681–701, 2019.

CHWIF, L.; MEDINA, A. C. **Modelagem e Simulação de Eventos Discretos**. 4. ed. São Paulo: ELSEVIER, 2015.

COSTA, A. F. B.; EPPRECHT, E. K.; CARPINETTI, L. C. R. **Controle Estatístico de Qualidade**. São Paulo: Atlas, 2018.

D'ANGELO, A.; CHONG, E. K. P. **A systems engineering approach to incorporating the Internet of Things to reliability-risk modeling for ranking conceptual designs**. Proceedings of International Mechanical Engineering Congress and Exposition 2018. **Anais...**Pittsburgh, PA: 2018.

DONHAUSER, T.; EBERSBACH, T.; FRANKE, J.; SCHUDERER, P. Rolling-reactive Optimization of Production in a Calcium Silicate Masonry Unit Plant Using Online Simulation. **Procedia CIRP**, v. 72, p. 249–254, 2018.

ELNIMR, A.; FAGIAR, M.; MOHAMED, Y. Two-way integration of 3D visualization and discrete event simulation for modeling mobile crane movement under dynamically changing site layout. **Automation in Construction**, v. 68, p. 245–248, 2016.

EYRE, J.; SCOTT, R.; FREEMAN, C. **Demonstration of an Industrial Framework for an Implementation of a Process Digital Twin**. Proceedings of International Mechanical Engineering Congress and Exposition 2018. **Anais...**Pittsburgh, PA: 2018.

FISHMAN, G. S. **Discrete-Event Simulation: modeling, programming and analysis**. 1. ed. New York: Springer, 2001.

GABRIEL, G. T.; CAMPOS, A.; LEAL, F.; MONTEVECHI, J. A. B. Good practices and

deficiencies in conceptual modelling: A systematic literature review. **Journal of Simulation**, v. 16, p. 84–100, 2022.

GHOSH, A. K. On optimum choice of k in nearest neighbor classification. **Computational Statistics & Data Analysis**, v. 50, p. 3113–3123, 2006.

GIL, A. C. **Como elaborar projetos de pesquisa**. 1. ed. São Paulo: Atlas, 2002.

GOLDSMAN, D.; NANCE, R. E.; WILSON, J. R. **A brief history of simulation revisited**. Proceedings of the 2010 Winter Simulation Conference. **Anais...IEEE**, 2010.

GONG, C. A.; SU, C. S.; TSENG, K. Implementation of Machine Learning for Fault Classification on Vehicle Power Transmission System. **IEEE Sensors Journal**, v. 20, n. 24, p. 1–14, 2020.

GOODALL, P.; SHARPE, R.; WEST, A. A data-driven simulation to support remanufacturing operations. **Computers in Industry**, v. 105, p. 48–60, 2019.

GREASLEY, A.; OWEN, C. Modelling people's behaviour using discrete-event simulation: a review. **International Journal of Operations & Production Management**, v. 38, n. 5, p. 1228–1244, 2018.

GRUBE, D.; MALIK, A. A.; BILBERG, A. SMEs can touch Industry 4.0 in the Smart Learning Factory. **Procedia Manufacturing**, v. 31, p. 219–224, 2019.

GUPTA, A. K.; SIVAKUMAR, A. I. Conjunctive Simulated Scheduling. **International Journal of Advanced Manufacturing Technology**, v. 26, p. 1409–1413, 2005.

HARPER, A.; MUSTAFEE, N. **A Hybrid Modelling Approach Using Forecasting and Real-Time Simulation to Prevent Emergency Department Overcrowding**. Proceedings of the 2019 Winter Simulation Conference. **Anais...Maryland**: 2019a.

HARPER, A.; MUSTAFEE, N. **Proactive Service Recovery in Emergency Departments: A Hybrid Modelling Approach using Forecasting and Real-Time Simulation**. Proceedings of ACM SIGSIM Conference 2019. **Anais...Chicago, USA**: 2019b.

HART, M. K.; HART, R. F.; SCHMALTZ, S. Control Limits for p Control Charts With Small Subgroup Sizes. **Quality Management in Health Care**, v. 16, n. 2, p. 123–129, 2007.

HASSANAT, A. B.; ABBADI, M. A.; ALHASANAT, A. A. Solving the Problem of the K Parameter in the KNN Classifier Using an Ensemble Learning Approach. **International Journal of Computer Science and Information Security**, v. 12, n. 8, p. 33–39, 2014.

JAHANGIRIAN, M.; EL DABI, T.; NASEER, A.; STERGILOULAS, L.; YOUNG, T. Simulation in manufacturing and business: A review. **European Journal of Operational Research**, v. 203, n. 1, p. 1–13, 2010.

JUNIOR, F. J. H.; BABIN, B.; MONEY, H.; SAMOUEL, P. **Fundamentos de métodos de pesquisa em administração**. 1. ed. São Paulo: Bookman, 2005.

KARAKRA, A. FONTANILI, F.; LAMINE, E.; LAMOTHE, J.; TAWHEEL, A. **Pervasive computing integrated Discrete Event Simulation for a hospital Digital Twin**. Proceedings of International Conference on Computer Systems and Applications 2018. *Anais...Aqaba, Jordan*: 2018.

KATZ, D.; MANIVANNAN, S. **Exception management on a shop floor using Online Simulation**. Proceedings of the 1993 Winter Simulation Conference. *Anais...Los Angeles, CA*: 1993.

KHALEGHI, A. M.; XU, D.; WANG, Z.; LI, M.; LOBOS, A.; LIU, J.; SON, Y. A DDDAMS-based planning and control framework for surveillance and crowd control via UAVs and UGVs. *Expert Systems with Applications*, v. 40, p. 7168–7183, 2013.

KIDDER, L. H. **Métodos de pesquisa nas relações sociais**. 4. ed. São Paulo: Editora Pedagógica e Universitária Ltda., 2004.

KITAZAWA, M.; TAKAHASHI, S.; TAKAHASHI, T.; YOUSHIKAWA, A.; TERANO, T. **Combining Workers' Behavior Data and Real Time Simulator for a Cellular Manufacturing System**. Proceedings of 2016 World Automation Congress (WAC). *Anais...Rio Grande, Puerto Rico*: 2016a.

KITAZAWA, M.; TAKAHASHI, S.; TAKAHASHI, T.; YOUSHIKAWA, A.; TERANO, T. **Improving a Cellular Manufacturing System through Real Time-Simulation and - Measurement**. Proceedings of Annual Computer Software and Applications Conference 2016. *Anais...Atlanta, GA*: 2016b.

KRITZINGER, W.; KARNER, M.; TRAAR, G.; HENJES, J.; SIHN, W. Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine*, v. 51, n. 11, p. 1016–1022, 2018.

KUMAR, R. P.; MELCHER, D.; BUTTOLO, P.; JIA, Y. Monitoring of occupant states in autonomous vehicles using capacitance-sensing imaging. *IEEE Sensors Journal*, v. 20, n. 21, p. 12914–12927, 2020.

KUNATH, M.; WINKLER, H. **Integrating the Digital Twin of the manufacturing system into a decision support system for improving the order management process**. 51st CIRP Conference on Manufacturing Systems Integrating. *Anais...Elsevier B.V.*, 2018.

LAW, A. M. **How to build valid and credible simulation models**. Proceedings of the 2009 Winter Simulation Conference M. *Anais...Austin, TX*: 2009.

LAW, A. M. **Simulation Modeling and Analysis**. 5. ed. Boston: McGraw-Hill Science, 2014.

LAZAR, J.; FENG, J. H.; HOCHHEIZER, H. **Research Methods in Human Computer Interaction**. 2. ed. New York: John Wiley & Sons Inc, 2017.

LEAL, F. et al. A practical guide for operational validation of discrete simulation models. *Pesquisa Operacional*, v. 31, n. 1, p. 57–77, 2011.

LEE, C.; HUANG, K.; SHEN, Y.; LEE, Y. Improved Weighted k -Nearest Neighbor Based on PSO for Wind Power System State Recognition. **Energies**, v. 13, p. 1–16, 2020.

LEE, H. Real-time manufacturing modeling and simulation framework using augmented reality and stochastic network analysis. **Virtual Reality**, v. 23, n. 1, p. 85–99, 2019.

LEE, J.; BAGHERI, B.; KAO, H. A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. **Manufacturing Letters**, v. 3, p. 18–23, 2015.

LEE, S.; RAMAKRISHNAN, S.; WYSK, R. A. **A federation Object Coordinator for Simulation Based Control and Analysis**. Proceedings of the 2002 Winter Simulation Conference. **Anais...**San Diego, CA: 2002.

LENG, J.; ZHANG, H.; YAN, D.; LIU, Q.; CHEN, X.; ZHANG, D. Digital Twin-Driven Manufacturing Cyber-Physical System for Parallel Controlling of Smart Workshop. **Journal of Ambient Intelligence and Humanized Computing**, v. 10, n. 3, p. 1155–1166, 2019.

LENG, J.; LIU, Q.; YE, S.; JUNG, J.; WANG, Y.; ZHANG, C.; ZHANG, D.; CHEN, X. Digital twin-driven rapid reconfiguration of the automated manufacturing system via an open architecture model. **Robotics and Computer-Integrated Manufacturing**, v. 63, p. 1–12, 2020.

LENHARD, J. Computer simulation: The cooperation between experimenting and modeling. **Philosophy of Science**, v. 74, n. 2, p. 176–194, 2007.

LEONI, R. C.; COSTA, A. F. B. The Shewhart attribute chart with alternated charting statistics to monitor bivariate and trivariate mean vectors. **Computers & Industrial Engineering**, v. 122, p. 273–282, 2018.

LI, G.; TAN, M. H.; NG, S. H. Metamodel-Based Optimization of Stochastic Computer Models for Engineering Design under Uncertain Objective Function. **IIE Transactions**, v. 51, n. 5, p. 517–530, 2019.

LILIS, G.; CUTSEM, O. VAN; KAYAL, M. **Building Virtualization Engine: a Novel Approach Based on Discrete Event Simulation**. Proceedings of International Conference on Event-based Control, Communication, and Signal Processing 2016. **Anais...**2016.

LILIS, G.; VAN CUTSEM, O.; KAYAL, M. A High-Speed Integrated building emulation engine based on discrete event simulation. **Journal of Systems Architecture**, v. 92, p. 53–65, 2019.

LIN, H.; SAMBAMOORTHY, S.; SHUKLA, S.; THORP, J.; MILI, L. **Power system and communication network co-simulation for smart grid applications**. Proceedings IEEE PES Innovative Smart Grid Technologies Conference Europe 2011. **Anais...**Manchester, UK: 2011.

LIN, H.; VEDA, S.; SHUKLA, S.; MILI, L.; THORP, J. GECO: Global event-driven co-simulation framework for interconnected power system and communication network. **IEEE Transactions on Smart Grid**, v. 3, n. 3, p. 1444–1456, 2012.

LIU, Y.; XU, X. Industry 4.0 and Cloud Manufacturing: A Comparative Analysis. **Journal of Manufacturing Science and Engineering**, v. 139, p. 1–8, 2017.

LIU, Q.; ZHANG, H.; LENG, J.; CHEN, X. Digital Twin-Driven Rapid Individualised Designing of Automated Flow-Shop Manufacturing System. **International Journal of Production Research**. v. 57, n. 12, p. 3903–3919, 2019.

LIU, Q.; LENG, J.; YAN, D. ZHANG, D.; WEI, L.; YU, A.; ZHAO, R.; ZHANG, H.; CHEN, X. Digital Twin-Based Designing of the Configuration, Motion, Control, and Optimization Model of a Flow-Type Smart Manufacturing System. **Journal of Manufacturing Systems**, V. 58, p. 52 – 64, 2021.

LONGO, F.; NICOLETTI, L.; PADOVANO, A. Smart operators in industry 4.0: A human-centered approach to enhance operators' capabilities and competencies within the new smart factory context. **Computers and Industrial Engineering**, v. 113, p. 144–159, 2017.

LOPES, M. A.; ALMEIDA, Á. S.; ALMADA-LOBO, B. Forecasting the medical workforce: a stochastic agent-based simulation approach. **Health Care Management Science**, v. 21, n. 1, p. 52–75, 2018.

LOPES, M. R.; COSTIGLIOLA, A.; PINTO, R.; VIEIRA, S.; SOUSA, J. Pharmaceutical quality control laboratory digital twin – A novel governance model for resource planning and scheduling. **International Journal of Production Research**, v. 7543, p. 1–15, 2019.

LU, Y. Industry 4.0: A survey on technologies, applications and open research issues. **Journal of Industrial Information Integration**, v. 6, p. 1–10, 2017.

LU, Y.; MIN, Q.; LIU, Z.; WANG, Y. An IoT-enabled simulation approach for process planning and analysis: a case from engine re- manufacturing industry. **International Journal of Computer Integrated Manufacturing**, v. 32, n. 4, p. 413–429, 2019.

MACAL, C.; NORTH, M. Introductory tutorial: **Agent-based modeling and simulation**. Proceedings of the 2014 Winter Simulation Conference. **Anais...**2014

MAHDAVI, I.; SHIRAZI, B.; SOLIMANPUR, M. Development of a simulation-based decision support system for controlling stochastic flexible job shop manufacturing systems. **Simulation Modelling Practice and Theory**, v. 18, p. 768–786, 2010.

MANAVALAN, E.; JAYAKRISHNA, K. A review of Internet of Things (IoT) embedded sustainable supply chain for industry 4.0 requirements. **Computers & Industrial Engineering**, v. 127, p. 925–953, 2019.

MENG, C.; NAGESHWARANIYER, S.; MAGHSOUDI, A.; SON, Y.; DESSEUREAULT, S. Data-driven modeling and simulation framework for material handling systems in coal mines. **Computers and Industrial Engineering**, v. 64, n. 3, p. 766–779, 2013.

MIETH, C.; MEYER, A.; HENKE, M. Framework for to the usage of data from real-time indoor localization systems derive for manufacturing simulation. **Procedia CIRP**, v. 81, p. 868–873, 2019.

MIGUEL, P. A. C.; FLEURY, A.; MELLO, C. H. P.; NAKANO, D. N.; LIMA, E. P.; TURRIONI, J. B.; HO, L. L.; MORABITO, R.; COSTA, S. E. G.; MARTINS, R. A.; SOUSA, R.; PUREZA, V. **Metodologia de pesquisa em Engenharia de Produção e Gestão de operações**. 3. ed. Rio de Janeiro: Elsevier, 2018.

MOALLEMI, M.; WAINER, G. **Designing an interface for real-time and embedded DEVS**. Proceedings of Spring Simulation Multiconference 2010. **Anais...**Orlando, FL: 2010

MOEUF, A.; PELLERIN, R.; LAMOURI, S.; TAMAYO, S.; BARBARAY, R. The industrial management of SMEs in the era of Industry 4.0. **International Journal of Production Research**, v. 56, n. 3, p. 1118–1136, 2018.

MONTEVECHI, J. A. B.; PEREIRA, T. F.; SILVA, C. E. S.; Miranda, R. C.; SCHEIDEGGER, A. P. G. **Identification of the main methods used in simulation projects**. Proceedings of the 2015 Winter Simulation Conference. **Anais...**Huntington Beach, CA: IEEE, 2015

MONTEVECHI, J. A. B.; SANTOS, C.; GABRIEL, G.; OLIVEIRA, M.; QUEIROZ, J.; LEAL, F. **A Method Proposal for Conducting Simulation Projects in Industry 4.0: a Cyber-Physical System in an Aeronautical Industry**. Proceeding of the 2020 Winter Simulation Conference. **Anais...**Orlando, USA: 2020

MONTGOMERY, D. C. **Introduction to Statistical Quality Control**. [s.l: s.n.]. v. 10

MOURTZIS, D. Simulation in the design and operation of manufacturing systems : state of the art and new trends. **International Journal of Production Research**, v. 58, n. 7, p. 1927–1949, 2020.

MUELLER, E.; CHEN, X.; RIEDEL, R. Challenges and Requirements for the Application of Industry 4.0: A Special Insight with the Usage of Cyber-Physical System. **Chinese Journal of Mechanical Engineering**, v. 30, n. 5, p. 1050–1057, 2017.

MURPHY, A.; TAYLOR, C.; ACHESON, C.; BUTTERFIELD, J.; JIN, Y.; HIGGINS, P.; COLLINS, R.; HIGGINS, C. Representing financial data streams in digital simulations to support data flow design for a future Digital Twin. **Robotics and Computer Integrated Manufacturing**, v. 61, 2020.

NANCE, R. E.; SARGENT, R. G. Perspectives on the Evolution of Simulation. **Operations Research Publication**, v. 50, n. 1, p. 161–172, 2002.

NEGAHBAN, A.; SMITH, J. S. Simulation for manufacturing system design and operation: Literature review and analysis. **Journal of Manufacturing Systems**, v. 33, p. 241–261, 2014.

NI, D.; XIAO, Z.; LIM, M. K. A systematic review of the research trends of machine learning in supply chain management. **International Journal of Machine Learning and Cybernetics**, v. 11, n. 7, p. 1463–1482, 2020.

NICOLETTI, L.; PADOVANO, A. Human factors in occupational health and safety 4.0: A cross-sectional correlation study of workload, stress and outcomes of an industrial emergency response. **International Journal of Simulation and Process Modelling**, v. 14, n. 2, p. 178–

195, 2019.

OLAITAN, O.; GERAGHTY, J.; YOUNG, P.; DAGKAKIS, G.; HEAVEY, C.; BAYER, M.; PERRIN, J.; ROBIN, S. Implementing ManPy, a Semantic-free Open-source Discrete Event Simulation Package, in a Job Shop. **Procedia CIRP**, v. 25, p. 253–260, 2014.

OLIVEIRA, J. B.; LIMA, R.; MONTEVECHI, J. A. B. Perspectives and relationships in Supply Chain Simulation: A systematic literature review. **Simulation Modelling Practice and Theory**, v. 62, p. 166–191, 2016.

ONGGO, B. S.; MUSTAFEE, N.; SMART, A.; JUAN, A.; MOLLOY, O. **Symbiotic Simulation System: Hybrid Systems Model Meets Big Data**. Proceedings of the 2018 Winter Simulation Conference. **Anais...Gothenburg, Sweden: 2018**

OPACIC, L.; SOWLATI, T. Applications of discrete-event simulation in the forest products sector: A review. **Forest Products Journal**, v. 67, n. 3–4, p. 219–229, 2017.

OSCARSSON, J.; MORIS, M. U. **Documentation of discrete event simulation models for manufacturing system life cycle simulation**. Proceedings of the 2002 Winter Simulation Conference. **Anais...2002**

PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, E. Scikit-learn: Machine Learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011.

PEREIRA, A. C.; ROMERO, F. A review of the meanings and the implications of the Industry 4.0 concept. **Procedia Manufacturing**, v. 13, p. 1206–1214, 2017.

PRAJAPAT, N. TIWARI, A.; TIWARI, D.; TURNER, C.; HUTABARAT, W. A **Framework for Next Generation Interactive and Immersive DES Models**. Proceedings of International Conference on Industrial Informatics 2019. **Anais...Helsinki, Finland: 2019**

PRINCES, E. Facing disruptive challenges in supply chain 4.0. **International Journal of Supply Chain Management**, v. 9, n. 4, p. 52–57, 2020.

PRIYADHARSHINI, G.; JOSEPHIN, J. S. F. A comprehensive review of various data collection approaches, features, and algorithms used for the classification of driving style. **IOP Conference Series: Materials Science and Engineering**, v. 993, p. 1–14, 2020.

PRODANOV, V. C. C.; FREITAS, E. C. **Metodologia do Trabalho Científico: métodos e técnicas da pesquisa e do trabalho acadêmico**. 2. ed. Novo Hamburgo: Universidade Feevale, 2013.

QI, Q.; TAO, F. Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison. **IEEE Access**, v. 6, p. 3585–3593, 2018.

RAJU, N. V. G.; LAKSHMI, K.; JAIN, V.; KALIDINDI, A.; PADMA, V. **Study the Influence of Normalization / Transformation process on the Accuracy of Supervised**

Classification. Proceedings of the 2020 International Conference on Smart Systems and Inventive Technology. **Anais...2020**

RENIERS, M. A.; MORTEL-FRONCZAK, J. M. VA. DE. An Engineering Perspective on Model-Based Design of Supervisors. **IFAC-PapersOnLine**, v. 51, n. 7, p. 257–264, 2018.

ROBINSON, S. Conceptual modelling for simulation Part II: A framework for conceptual modelling. **Journal of the Operational Research Society**, v. 59, n. 3, p. 291–304, 2008.

RODIČ, B. Industry 4.0 and the New Simulation Modelling Paradigm. **Organizacija**, v. 50, n. 3, p. 193–207, 2017.

RODIČ, B.; KANDUČ, T. Optimisation of a complex manufacturing process using discrete event simulation and a novel heuristic algorithm. **International Journal of Mathematical Models and Methods in Applied Sciences**, v. 9, p. 320–329, 2015.

ROSSMANN, J.; SCHLUSE, M.; WASPE, R. **3D Discrete Event Systems: an efficient way to model and supervise dynamic behavior in virtual environments.** Proceedings of International Design Engineering Technical Conferences & Computers and Information in Engineering Conference 2008. **Anais...New York, USA: 2008**

RÜTTIMANN, B. G.; STÖCKLI, M. T. Lean and Industry 4.0—Twins, Partners, or Contenders? A Due Clarification Regarding the Supposed Clash of Two Production Systems. **Journal of Service Science and Management**, v. 09, n. 06, p. 485–500, 2016.

SABAHNO, H.; CASTAGLIOLA, P.; AMIRI, A. A variable parameters multivariate control chart for simultaneous monitoring of the process mean and variability with measurement errors. **Quality and Reliability Engineering International**, v. 36, n. 4, p. 1161–1196, 2020.

SAEZ, M.; MATURANA, F.; BARTON, K.; TILBURY, D. Real-Time Manufacturing Machine and System Performance Monitoring Using Internet of Things. **IEEE Transactions on Automation Science and Engineering**, v. 15, n. 4, p. 1735–1748, 2018.

SALAMA, S.; ELTAWIL, A. B. A Decision Support System Architecture Based on Simulation Optimization for Cyber-Physical Systems. **Procedia Manufacturing**, v. 26, p. 1147–1158, 2018.

SANTOS, C. H.; GABRIEL, G.; AMARAL, J.; MONTEVECHI, J. A. B.; QUEIROZ, J. Decision-making in a fast fashion company in the Industry 4.0 era: a Digital Twin proposal to support operational planning. **The International Journal of Advanced Manufacturing Technology**, v. 116, n. 5, p. 1653–1666, 2021.

SANTOS, C. H.; QUEIROZ, J.; LEAL, F.; MONTEVECHI, J. Use of Simulation in the Industry 4.0 context: creation of a Digital Twin to optimize decision making on non-automated process. **Journal of Simulation**, v. 16, n. 3, p. 284–297, 2022a.

SANTOS, C. H.; MONTEVECHI, J. A. B.; QUEIROZ, J.; MIRANDA, R. C.; LEAL, F. Decision support in productive processes through DES and ABS in the Digital Twin era: a systematic literature review. **International Journal of Production Research**, v. 60, n. 8, p. 2662–2681, 2022b.

SANTOS, C. H.; CAMPOS, A. T.; MONTEVECHI, J. A.; MIRANDA, R. C.; COSTA, A. Digital Twin simulation models: a validation method based on Machine Learning and Control Charts. **International Journal of Production Research**, v. x, p. 1-18, 2023.

SARGENT, R. G. **Verification and Validation of Simulation Models: An Advanced Tutorial**. Proceedings of the 2011 Winter Simulation Conference. **Anais...** Phoenix, Arizona: 2011.

SARGENT, R. G. Verification and validation of simulation models. **Journal of Simulation**, v. 7, n. 1, p. 12–24, 2013.

SARGENT, R. G. An interval statistical procedure for use in validation of simulation models. **Journal of Simulation**, v. 9, p. 1–6, 2014.

SARGENT, R. G.; GOLDSMAN, D. M.; YAACOUB, T. **A tutorial on the operational validation of simulation models**. Proceedings of the 2016 Winter Simulation Conference. **Anais...**Arlington, Virginia: 2016.

SCHEIDEGGER, A. P. G.; PEREIRA, T.; OLIVEIRA, M.; BANERJEE, A.; MONTEVECHI, J. A. An introductory guide for hybrid simulation modelers on the primary simulation methods in industrial engineering identified through a systematic review of the literature. **Computers & Industrial Engineering**, v. 124, p. 474–492, 2018.

SCHOLL, W.; GAN, B.; PEH, M.; LENDERMANN, P.; NOARK, D.; ROSE, O.; PREUSS, P. **Towards realization of a high-fidelity simulation model for short-term horizon forecasting in wafer fabrication facilities**. Proceedings of the 2010 Winter Simulation Conference. **Anais...**Baltimore, USA: 2010.

SCHOLL, W.; MOSINSKI, M.; GAN, B.; LENDERMANN, P.; PREUSS, P.; NOACK, D. **A multi-stage Discrete Event Simulation approach for scheduling of maintenance activities in a semiconductor manufacturing line**. Proceedings of the 2012 Winter Simulation Conference. **Anais...**Berlin, Germany: 2012.

SEMERARO, C.; LEZOUCHE, M.; PANETTO, H.; DASSISTI, M. Digital twin paradigm: A systematic literature review. **Computers in Industry**, v. 130, p. 1–23, 2021.

SHAFTO, M.; CONROY, M.; DOYLE, R.; GLAESSGEN, E.; KLEMP, C.; JACQUELINE, L.; WANG, L. DRAFT Modeling, Simulation, Information Technology & Processing Roadmap. **Technology Area 11 - National Aeronautics and Space Administration (NASA)**, p. 1–27, 2010.

SHANG, H.; WAINER, G. A. Dynamic Structure DEVS: **Improving the Real-Time Embedded Systems Simulation and Design**. Proceedings of 41st Annual Simulation Symposium Dynamic. **Anais...**Ottawa, Canada: 2008.

SHARMA, R.; KAMBLE, S.; GUNASEKARAN, A.; KUMAR, V.; KUMAR, A. A systematic literature review on machine learning applications for sustainable agriculture supply chain performance. **Computers and Operations Research**, v. 119, p. 104926, 2020.

SHEWHART, W. A. Quality Control Charts. **Bell Labs Technical Journal**, v. 5, n. 4, p. 593-

603, 1926.

SHIN, H.; CHO, K.-W.; OH, C.-H. SVM-Based Dynamic Reconfiguration CPS for Manufacturing System in Industry 4.0. **Wireless Communications and Mobile Computing**, p. 1–15, 2018.

SHIRAZI, B.; MAHDAVI, I.; SOLIMANPUR, M. Development of a Simulation-Based Intelligent Decision Support System for the Adaptive Real-Time Control of Flexible Manufacturing Systems. **Journal of Software Engineering & Applications**, v. 3, p. 661–673, 2010.

SICKEL, J. H. V; LEE, K. Y. **Real-time Based Agent Architecture for Power Plant Control**. Proceedings International Conference on Intelligent System Applications to Power Systems 2019. **Anais...**Curitiba, Brazil: 2009.

SIEBERS, P. O.; MACAL, C.; GARNETT, J.; BUXTON, D.; PIDD, M. Discrete-event simulation is dead, long live agent-based simulation! **Journal of Simulation**, v. 4, p. 204–210, 2010.

SIMMONDS, R.; BRADFORD, R.; UNGER, B. **Applying Parallel Discrete Event Simulation to Network Emulation**. Proceedings of Fourteenth Workshop on Parallel and Distributed Simulation 2000. **Anais...**Bologna, Italy: 2014.

SIVAKUMAR, A. I.; CHONG, C. S. **Development of an Object Oriented Simulation Engine for On-Line Simulation and Optimization**. Proceedings 33rd Annual Simulation Symposium 2000. **Anais...**Washington, DC: 2000.

SIVAKUMAR, A. I.; GUPTA, A. K. Online Multiobjective Pareto Optimal Dynamic Scheduling of Semiconductor Back-End Using Conjunctive Simulated Scheduling. **IEEE Transactions on Electronics Packaging Manufacturing**, v. 29, n. 2, p. 99–109, 2006.

SKOOGH, A.; PERERA, T.; JOHANSSON, B. Input data management in simulation - Industrial practices and future trends. **Simulation Modelling Practice and Theory**, v. 29, p. 181–192, 2012.

SOLHMIRZAEI, R.; SALEHI, H.; KODUR, V.; NASER, M. Machine learning framework for predicting failure mode and shear capacity of ultra high performance concrete beams. **Engineering Structures**, v. 224, p. 1–11, 2020.

SONG, L.; EL-DIN, N. N. Adaptive real-time tracking and simulation of heavy construction operations for look-ahead scheduling. **Automation in Construction**, v. 27, p. 32–39, 2012.

SONG, L.; RAMOS, F.; ARNOLD, K. **A framework for real-time simulation of heavy construction operations**. Proceedings of the 2008 Winter Simulation Conference. **Anais...**Miami, FL: 2008.

SONY, M. Industry 4.0 and lean management: a proposed integration model and research propositions. **Production & Manufacturing Research**, v. 6, n. 1, p. 416–432, 2018.

SORMAZ, D. N.; MALIK, M. Data-driven Simulation Modelling for Progressive Care Units

in Hospitals. **Procedia Manufacturing**, v. 17, p. 819–826, 2018.

STERINGER, R.; ZORRER, H.; ZAMBAL, S. EITZINGER, C. Using Discrete Event Simulation in Multiple System Life Cycles to support Zero-Defect Composite Manufacturing in Aerospace Industry. **IFAC PapersOnLine**, v. 52, n. 13, p. 1467–1472, 2019.

TAKO, A. A.; ROBINSON, S. Model development in discrete-event simulation and system dynamics: An empirical study of expert modellers. **European Journal of Operational Research**, v. 207, n. 2, p. 784–794, 2010.

TAO, F.; CHENG, J.; QI, Q.; ZHANG, M.; ZHANG, H.; SUI, F. Digital twin-driven product design, manufacturing and service with big data. **International Journal of Advanced Manufacturing Technology**, v. 94, p. 3563–3576, 2018.

TAO, F.; ZHANG, M. Digital Twin Shop-Floor: A New Shop-Floor Paradigm Towards Smart Manufacturing. **IEEE Access**, v. 5, p. 20418–20427, 2017.

TERKAJ, W.; GABORADI, P.; TREVISAN, C.; TOLIO, T.; URGO, M. A digital factory platform for the design of roll shop plants. **CIRP Journal of Manufacturing Science and Technology**, v. 26, p. 88–93, 2019.

THANOS, A. E.; BASTANI, M.; CELIK, N.; CHEN, C. Framework for Automated Control in Microgrids. **IEEE Transactions on Smart Grid**, v. 8, n. 1, p. 1–10, 2017.

TIACCI, L. Object-Oriented Event-Graph modeling formalism to simulate manufacturing systems in the Industry 4.0 era. **Simulation Modelling Practice and Theory**, v. 99, p. 1–33, 2020.

TORTORELLA, G. L.; FETTERMANN, D. Implementation of industry 4.0 and lean production in brazilian manufacturing companies. **International Journal of Production Research**, v. 56, n. 8, p. 2975–2987, 2018.

TRANFIELD, D.; DENYER, D.; SMART, P. Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. **British Journal of Management**, v. 14, p. 207–222, 2003.

TRAPPEY, A. J. C.; TRAPPEY, C.; GOVINDARAJAN, U.; CHUANG, A.; SUN, J. A review of essential standards and patent landscapes for the Internet of Things: A key enabler for Industry 4.0. **Advanced Engineering Informatics**, v. 33, p. 208–229, 2017.

TRIEBIG, C.; KLÜGL, F. Elements of a documentation framework for agent-based simulation models. **Cybernetics and Systems**, v. 40, n. 5, p. 441–474, 2009.

TRIGUEIRO, W. DE S. J. MONTEVECHI, J.; MIRANDA, R.; CAMPOS, A. Discrete simulation-based optimization methods for industrial engineering problems: A systematic literature review. **Computers & Industrial Engineering**, v. 128, p. 526–540, 2019.

UHLEMANN, T. H.; LEHMANN, C.; STEINHILPER, R. The Digital Twin: Realizing the Cyber-Physical Production System for Industry 4.0. **Procedia CIRP**, v. 61, p. 335–340, 2017.

- URIARTE, A. G.; NG, A. H. C.; MORIS, M. U. Supporting the lean journey with simulation and optimization in the context of Industry 4.0. **Procedia Manufacturing**, v. 25, p. 586–593, 2018.
- VAHDATIKHAKI, F.; HAMMAD, A. Framework for near real-time simulation of earthmoving projects using location tracking technologies. **Automation in Construction**, v. 42, p. 50–67, 2014.
- VAHDATIKHAKI, F.; HAMMAD, A. Risk-based look-ahead workspace generation for earthwork equipment using near real-time simulation. **Automation in Construction**, v. 58, p. 207–220, 2015.
- VIJAYAKUMAR, K.; DHANASEKARAN, C.; PUGAZHENTHI, R.; SIVAGANESAN, S. Digital Twin for Factory System Simulation. **International Journal of Recent Technology and Engineering**, v. 8, n. 1, 2019.
- WAN, Y.; CAO, J.; ZHANG, S.; TU, G.; LU, C.; XU, X.; LI, K. An integrated cyber-physical simulation environment for smart grid applications. **Tsinghua Science and Technology**, v. 19, n. 2, p. 133–143, 2014.
- WANG, J.; QIANCHUAN, Z.; YIN, Z. **An effective framework to simulate the cyber-physical systems with application to the building and energy saving**. Proceedings Chinese Control Conference 2013. **Anais...Xi'an, China: 2013**.
- WANG, M. C.; TSAI, C. F.; LIN, W. C. Towards missing electric power data imputation for energy management systems. **Expert Systems with Applications**, v. 174, p. 1–9, 2021.
- WANG, Z. **Selecting verification and validation techniques for simulation projects: A planning and tailoring strategy**. Proceedings of the 2013 Winter Simulation Conference. **Anais...IEEE, 2013**.
- WANG, Z.; LI, M.; KHALEGHI, A.; XU, D.; LOBOS, A.; VO, C.; LIEN, J.; LIU, J.; SON, Y. DDDAMS-based Crowd Control via UAVs and UGVs. **Procedia Computer Science**, v. 18, p. 2028–2035, 2013.
- WANG, Z.; LIAO, X.; WANG, C.; OSWALD, D.; WU, G.; BORIBOONSOMSIN. Driver Behavior Modeling using Game Engine and Real Vehicle: A Learning-Based Approach. **IEEE Transactions on Intelligent Vehicles**, v. 8858, p. 1–12, 2020.
- WINKELHAUS, S.; GROSSE, E. H. Logistics 4.0: a systematic review towards a new logistics system. **International Journal of Production Research**, v. 58, n. 1, p. 18–43, 2020.
- WRIGHT, L.; DAVIDSON, S. How to tell the difference between a model and a digital twin. **Advanced Modeling and Simulation in Engineering Sciences**, v. 7, n. 13, p. 1–13, 2020.
- XIAOBO, X.; KANGFENG, Z.; YIXIAN, Y.; GUOAI, X. **A Model for Real-time Simulation of Large-scale Networks based on Network Processor**. Proceedings of International Conference on Broadband Network & Multimedia Technology. **Anais...Beijing, China: 2009**.

XU, L. DA; XU, E. L.; LI, L. Industry 4.0: State of the art and future trends. **International Journal of Production Research**, v. 56, n. 8, p. 2941–2962, 2018.

YIN, Y.; STECKE, K. E.; LI, D. The evolution of production systems from Industry 2.0 through Industry 4.0. **International Journal of Production Research**, v. 56, p. 848–861, 2018.

ZHANG, J.; LIAN, J.; YI, Z.; YANG, S.; SHAN, Y. High-Accuracy Guide Star Catalogue Generation with a Machine Learning Classification Algorithm. **Sensors**, v. 21, p. 1–16, 2021.

ZHENG, X.; JULIEN, C.; CHEN, H.; PODOROZHNY, R.; CASSEZ, F. Real-time simulation support for runtime verification of cyber-physical systems. **ACM Transactions on Embedded Computing Systems**, v. 16, n. 4, p. 1–24, 2017.

ZHONG, R. Y.; XU, X.; KLOTZ, E.; NEWMAN, S. Intelligent Manufacturing in the Context of Industry 4.0: A Review. **Engineering**, v. 3, n. 5, p. 616–630, 2017.

ZHOU, W.; CHENG, C.; ZHENG, Z. Optimal design of an attribute control chart for monitoring the mean of autocorrelated processes. **Computers & Industrial Engineering**, v. 137, p. 1–15, 2019.

ZHOU, Y.; HUANG, K.; HU, D. **C2 Decision Support System Based on Symbiotic Simulation**. Proceedings of International Conference on System Simulation and Scientific Computing 2008. **Anais...Beijing, China: 2008**.

ZHUANG, C.; LIU, J.; XIONG, H. Digital twin-based smart production management and control framework for the complex product assembly shop-floor. **The International Journal of Advanced Manufacturing Technology**, v. 96, p. 1149–1163, 2018.

ZÖRRER, H.; STERINGER, R.; ZAMBAL, S.; EITZINGER, C. Using Business Analytics for Decision Support in Zero Defect Manufacturing of Composite Parts in the Aerospace Industry. **IFAC Papers OnLine**, v. 52, n. 13, p. 1461–1466, 2019.

ZUPAN, H.; JANEZ, Ž.; HERAKOVIČ, N. **Local Search with Discrete Event Simulation for the Job Shop Scheduling Problem**. Proceedings of SOHOMA'17 Workshop on Service Orientation in Holonic and Multi-Agent Manufacturing. **Anais...Carquefou, França: 2018**.

ZWETSLOOT, I. M.; WOODALL, W. H. A Review of Some Sampling and Aggregation Strategies for Basic Statistical Process Monitoring. **Journal of Quality Technology**, v. 53, p. 1–16, 2021.