FERNANDA BARRETO DE ALMEIDA ROCHA MARIZ

Análise comparativa dos modelos *Drum-Buffer-Rope* e *Constant Work-In-Process* em um ambiente com montagem e produção contra pedido

São Paulo - SP

2018

FERNANDA BARRETO DE ALMEIDA ROCHA MARIZ

Análise comparativa dos modelos *Drum-Buffer-Rope* e *Constant Work-In-Process* em um ambiente com montagem e produção contra pedido

Tese apresentada à Escola Politécnica da Universidade de São Paulo para obtenção do título de Doutor em Ciências.

Área de concentração: Engenharia de Produção

Orientador: Prof. Dr. Marco Aurélio de Mesquita

São Paulo

2018

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Catalogação-na-publicação

Mariz, Fernanda Barreto de Almeida Rocha
Análise comparativa dos modelos Drum-Buffer-Rope e Constant Work In-Process em um ambiente com montagem e produção contra pedido / F. B. A. R. Mariz -- São Paulo, 2018.
139 p.

Tese (Doutorado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Produção.

1.Controle da produção 2.Administração da produção 3.Estoques I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Produção II.t.

Ao meu amor, Lawrence;
Aos meus pais, Dorgival e Aristotelina;
Aos meus sogros, Valdir (in memoriam) e Lenia.

AGRADECIMENTOS

Agradeço, primeiramente, ao meu Deus, que me concedeu o dom da vida e tem dia após dia me ajudado a aprender com todas as oportunidades e desafios que Ele me permite viver. Não tenho palavras para agradecer a tua bondade e fidelidade em minha vida.

Ao Prof. Marco Aurélio de Mesquita, que sempre esteve disponível para me ensinar e compartilhar todo o seu conhecimento. Cada orientação é única e eu sou grata pela oportunidade que tive de aprender com você ao longo desses três anos de doutorado. Muito obrigada pela oportunidade e por ter sido sempre compreensível com minha rotina de viagens.

Ao meu esposo Lawrence, obrigada por compartilhar comigo todos esses dias tão intensos que vivi. Você abraçou os meus sonhos e planejamos juntos essa fase em nossas vidas. Obrigada pela compreensão em cada ausência, nas muitas renúncias e com as mais diversas crises de estresse e ansiedade. Sua calmaria e leveza foram essenciais para a conclusão dessa etapa.

Aos meus pais, toda a minha gratidão e o meu amor por tudo que fizeram por mim. Vocês são meu porto seguro. Sempre me incentivaram a crescer profissionalmente, estão sempre ao meu lado, orando por mim, me aconselhando e me ajudando a passar por todas as fases da vida.

Aos meus sogros, Valdir (*in memoriam*) e Lenia, que sempre se preocuparam conosco. Constantemente oraram por mim e por todos que estavam ao meu redor. Compreenderam todas as minhas ausências e torceram por cada passo dado na realização dessa conquista acadêmica. Em nome desses casais, agradeço a todos os meus familiares (irmãos, avó, tios e cunhadas).

Aos professores Débora Pettri Ronconi e Fernando Bernardi de Souza pelas contribuições na banca de qualificação. Muito obrigada Prof. Fernando, por estar sempre disponível para compartilhar os seus conhecimentos sobre a Teoria das Restrições.

Ao professor David Custódio de Sena, minha gratidão por mesmo sem me conhecer ter dispensado diversas vezes o seu tempo para me ensinar a modelar no *software AnyLogic*.

A amiga Marianna Campos, em nome de quem agradeço a todos os amigos e colegas de Natal/RN que me incentivaram e acompanharam essa empreitada.

A amiga Bruna Carvalho, pela companhia no doutorado, por sempre me receber em sua casa com carinho e me ensinar tanto com as nossas conversas. Manteremos esse elo por muitos anos.

Aos demais amigos do PPGEP/Poli/USP, Hanna, Victória, Sérgio, Augusto e Luiz Felipe, a companhia de vocês tornou essa caminhada mais leve e divertida. Vocês são parte dos meus presentes do doutorado.

Obrigada a secretária do programa Lídia, pela presteza e gentileza em seu fazer diário, em nome de quem agradeço aos demais servidores do Departamento de Engenharia de Produção/USP.

“Lembre da minha ordem: ‘Seja forte e corajoso!
Não fique desanimado, nem tenha medo, porque eu,
o Senhor, seu Deus, estarei com você em qualquer
lugar para onde você for!’” (Josué 1:9)

RESUMO

MARIZ, Fernanda Barreto de Almeida Rocha. **Análise comparativa dos modelos Drum-Buffer-Rope e Constant Work-In-Process em um ambiente com montagem e produção contra pedido**. 2018. 139f. Tese (Doutorado em Engenharia de Produção) – Escola Politécnica, Universidade de São Paulo, São Paulo, 2018.

Os modelos de programação e controle *Constant Work-In-Process* (CONWIP) e *Drum-Buffer-Rope* (DBR) são mecanismos bastante conhecidos na literatura de Planejamento e Controle da Produção, especialmente, pela robustez e bom desempenho. Esses modelos já foram comparados, sobretudo, em ambientes do tipo *flow shop*. Há muitos trabalhos que apontam o DBR como superior, outros que sugerem equivalência entre ambos e, em menor proporção, casos onde o CONWIP foi melhor. Contudo, estudos comparativos em ambientes mais complexos, como sistemas com montagem, ainda são escassos na literatura. Diante do exposto, a presente tese visa comparar o CONWIP e o DBR em um ambiente de montagem (*assembly flow shop*) contra pedido. Para tanto, foram desenvolvidos dois modelos de simulação por eventos discretos e um experimento hierárquico-cruzado, com os seguintes fatores: modelo de programação e controle, posição do gargalo, severidade do gargalo e limite de *buffer*. Os mecanismos foram avaliados quanto aos indicadores tempo médio de fluxo, *lead time* médio, utilização do gargalo, percentual de ordens atrasadas e atraso médio. Os resultados demonstraram que, para o ambiente em análise, os modelos se mostraram estatisticamente semelhantes, havendo diferenças significativas apenas para a variável tempo médio de fluxo, o que ratifica e estende conclusões de estudos anteriores. Há evidências de que o CONWIP realiza um controle mais efetivo do estoque em processo e não é significativamente afetado pelas diferentes posições do gargalo. Para o CONWIP, os melhores resultados foram identificados quando a restrição estava no início do sistema, enquanto que para o DBR os melhores resultados foram quando o gargalo estava na última estação da linha de fabricação (operação anterior a montagem).

Palavras-chave: *Constant Work-In-Process. Drum-Buffer-Rope.* Montagem. *Make-to-order*.

ABSTRACT

MARIZ, Fernanda Barreto de Almeida Rocha. **Análise comparativa dos modelos Drum-Buffer-Rope e Constant Work-In-Process em um ambiente com montagem e produção contra pedido**. 2018. 139f. Tese (Doutorado em Engenharia de Produção) – Escola Politécnica, Universidade de São Paulo, São Paulo, 2018.

The Constant Work-In-Process (CONWIP) and Drum-Buffer-Rope (DBR) scheduling and control models are well known mechanisms in the Production Planning and Control literature, especially for robustness and good performance. These models have already been compared, especially, to flow shop environments. There are many papers that point to DBR as superior, others that suggest equivalence between both and, less often, cases were CONWIP was better. However, comparative studies in more complex environments, such as assembly systems, are still rare in the literature. Therefore, this thesis aims to compare CONWIP and DBR in a make-to-order assembly flow shop environment. For that, two models of discrete event simulation and a crossed and nested experiment were developed, with the following factors: scheduling and control model, bottleneck position, bottleneck severity and buffer limit. The mechanisms were evaluated with respect to the indicators mean flow time, mean lead time, bottleneck utilization, percentage tardy and mean tardiness. The results showed that, for the environment under analysis, the models were statistically similar, with significant differences only for the variable mean flow time, which ratifies and extends the conclusions from previous studies. There is evidence that CONWIP performs more effective control of order release and is not significantly affected by the different bottleneck positions. For the CONWIP, the best results were identified when the constraint was at the beginning of the system, while for the DBR the best results were when the bottleneck was at the last station of the manufacturing line (before assembly operation).

Keywords: Constant Work-In-Process. Drum-Buffer-Rope. Assembly. Make-to-order.

LISTA DE FIGURAS

Figura 1 - Estrutura da tese 25
Figura 2 - Estratégias de resposta à demanda 26
Figura 3 - Controle da liberação de ordens 30
Figura 4 – Processo de gerenciamento de pedidos 30
Figura 5 - Representação do Kanban 31
Figura 6 - Representação do sistema POLCA 33
Figura 7 - Proteção do *time buffer* à programação do gargalo 37
Figura 8 - Estrutura básica do DBR com recurso restritivo e montagem 38
Figura 9 - Programação do gargalo e da liberação dos materiais 40
Figura 10 - Representação da lógica do CONWIP em uma linha 46
Figura 11 - CONWIP em linhas com montagem 47
Figura 12 - Decisão de dimensionamento do WIP 49
Figura 13 - Etapas da pesquisa 61
Figura 14 - *Assembly flow shop* em estudo com diferentes posições da estação gargalo 63
Figura 15 - Configuração dos *loops* e limites de WIP no modelo CONWIP 66
Figura 16 - Modelo conceitual da simulação do sistema CONWIP 68
Figura 17 - Captura da tela principal do modelo de simulação do CONWIP 71
Figura 18 – Validação do CONWIP 73
Figura 19 - Configurações do DBR de acordo com a posição do gargalo no ambiente em estudo 75
Figura 20 – Lógica da programação e controle da liberação com o DBR com o gargalo em diferentes posições 77
Figura 21 - Modelo conceitual da simulação do sistema DBR 79
Figura 22 - Captura da tela principal do modelo de simulação do DBR 81
Figura 23 - Gráfico de Gantt da programação das ordens até o gargalo (M3) 83
Figura 24 - Análise inicial do impacto do nível de maxWIP na vazão do sistema 90
Figura 25 - Análise inicial do impacto do nível de maxWIP no tempo médio de fluxo do sistema 91
Figura 26 - Análise inicial do impacto do nível de maxWIP no *lead time* médio do sistema 91
Figura 27 - Análise inicial do efeito do nível do maxWIP (10 a 24) na vazão do sistema 92
Figura 28 - Análise inicial do efeito do nível do maxWIP (10 a 24) no *flow time* e *lead time* médios 92
Figura 29 - Análise da relação entre o maxWIP1 e o maxWIP2 quanto ao tempo médio de fluxo e o lead time médio 93
Figura 30 - Análise inicial do efeito do time buffer no *flow time* e *lead time* médios com o gargalo na máquina 3 96
Figura 31 - Análise inicial do efeito do time buffer no *flow time* e *lead time* médios com o gargalo na máquina 9 97
Figura 32 - Distribuição dos resultados dos indicadores utilização do gargalo, tempo médio de fluxo e *lead time* médio no CONWIP 99
Figura 33 - Distribuição dos resultados dos indicadores atraso médio e percentual de ordens atrasadas no CONWIP 99
Figura 34 - Efeito dos fatores na utilização do gargalo no modelo CONWIP 100
Figura 35 - Efeito dos níveis de maxWIP no *flow time* e *lead time* médios no modelo CONWIP 100
Figura 36 - Efeito dos níveis de maxWIP no percentual de ordens atrasadas e atraso médio no modelo CONWIP 101
Figura 37 - Efeito da posição do gargalo no *flow time* e *lead time* médios no modelo CONWIP 101
Figura 38 - Efeito da posição do gargalo no percentual de ordens atrasadas e atraso médio no modelo CONWIP 102
Figura 39 - Efeito da severidade do gargalo no *flow time* e *lead time* médios no modelo CONWIP 103

Figura 40 - Efeito da severidade do gargalo no percentual de ordens atrasadas e atraso médio no modelo CONWIP 103
Figura 41 - Distribuição dos resultados dos indicadores utilização do gargalo, tempo médio de fluxo e *lead time* médio no DBR 104
Figura 42 - Distribuição dos resultados dos indicadores atraso médio e percentual de ordens atrasadas no DBR 105
Figura 43 - Efeito dos fatores na utilização do gargalo no modelo DBR 105
Figura 44 - Efeito do tbp no *flow time* e *lead time* médios no modelo DBR 106
Figura 45 - Efeito do tbp no percentual de ordens atrasadas e atraso médio no modelo DBR 106
Figura 46 - Efeito da posição do gargalo no *flow time* e *lead time* médios no modelo DBR 107
Figura 47 - Efeito da posição do gargalo no percentual de ordens atrasadas e atraso médio no modelo DBR 107
Figura 48 - Efeito da severidade do gargalo no *flow time* e *lead time* médios no modelo DBR 108
Figura 49 - Efeito da severidade do gargalo no percentual de ordens atrasadas e atraso médio no modelo DBR 108
Figura 50 - Desempenho dos modelos CONWIP e DBR quanto a utilização do gargalo 112
Figura 51 - Desempenho dos modelos CONWIP e DBR quanto ao tempo médio de fluxo 112
Figura 52 - Desempenho dos modelos CONWIP e DBR quanto ao *lead time* médio 113
Figura 53 - Tempo médio de espera na fila de entrada com os modelos CONWIP e DBR 113
Figura 54 - Desempenho dos modelos CONWIP e DBR quanto ao atraso médio 114
Figura 55 - Desempenho dos modelos CONWIP e DBR quanto ao percentual de ordens atrasadas 114
Figura 56 - Blocos do *AnyLogic* empregados nos modelos de simulação 135
Figura 57 - Análise do regime transitório para a variável utilização (CONWIP) 136
Figura 58 - Análise do regime transitório para a variável utilização (DBR) 136
Figura 59 - Análise do regime transitório para a variável tempo de fluxo (CONWIP) 137
Figura 60 - Análise do regime transitório para a variável tempo de fluxo (DBR) 137
Figura 61 - Análise do regime transitório para a variável *lead time* médio 137
Figura 62 - Análise do regime transitório para a variável *lead time* médio (DBR) 138
Figura 63 - Análise do regime transitório para a variável percentual de ordens atrasadas (CONWIP) 138
Figura 64 - Análise do regime transitório para a variável percentual de ordens atrasadas (DBR) 138
Figura 65 - Análise do regime transitório para a variável atraso médio (CONWIP) 139
Figura 66 - Análise do regime transitório para a variável atraso médio (DBR) 139

LISTA DE QUADROS

Quadro 1 - Definição do problema de pesquisa 21
Quadro 2 - Estudos sobre os modelos CONWIP e DBR publicados até 16 de dezembro de 2018 23
Quadro 3 - Características das estratégias de resposta à demanda 27
Quadro 4 - Sistematização das abordagens para o dimensionamento dos cartões 51
Quadro 5 - Comparativo dos modelos DBR e CONWIP 55
Quadro 6 - Características dos estudos comparativos 56

LISTA DE TABELAS

Tabela 1 - Síntese das características do ambiente simulado 64
Tabela 2 - Exemplo de teste da verificação do DBR..... 83
Tabela 3 - Cálculo dos intervalos de confiança com base em 50 replicações do modelo CONWIP ... 88
Tabela 4 - Cálculo dos intervalos de confiança com base em 50 replicações do modelo DBR..... 89
Tabela 5 - Resumo do resultado de 50 replicações do CONWIP para *benchmark* dos *time buffers* no DBR..... 95
Tabela 6 - Constantes para dimensionamento dos *time buffers* em função da posição do gargalo..... 96
Tabela 7 - Síntese dos fatores e níveis dos experimentos 97
Tabela 8 - ANOVA da variável de resposta tempo médio de fluxo..... 110
Tabela 9 - ANOVA da variável de resposta lead time médio 110
Tabela 10 - ANOVA da variável de resposta atraso médio 111
Tabela 11 - ANOVA da variável de resposta percentual de atraso..... 111
Tabela 12 - Resultado da comparação dos modelos com teste t 115
Tabela 13 - Impacto da posição do gargalo nos modelos CONWIP e DBR 119

LISTA DE SIGLAS

ANOVA – *Analysis of variance*

ATO – *Assemble-to-order*

BLG – *Backlog* do gargalo

BM – *Buffer Management*

BoM – *Bill of Materials*

CODP – *Customer Order Decoupling Point*

CONWIP – *Constant Work-In-Process*

DBR – *Drum-Buffer-Rope*

DFC – *Dynamic Flow Control*

EDD – *Earliest Due Date*

ETO – *Engineer-to-order*

FCFS – *First Come, First Served*

FS – *Flow shop*

JS – *Job shop*

maxWIP – *Maximum Work-In-Process*

MRP – *Material Requirements Planning*

MTO – *Make-to-order*

MTS – *Make-to-stock*

OPT – *Optimized Production Technology*

PCP – Planejamento e Controle da Produção

POLCA – *Paired-cell Overlapping Loops of Cards with Authorization*

PRD – *Planned Release Date*

QRM – *Quick Response Manufacturing*

RD – *Release Date*

S-DBR – *Simplified Drum-Buffer-Rope*

STC – *Statistical Throughput Control*

tb – *Time buffer*

tbe – *Time buffer* da expedição

tbg – *Time buffer* do gargalo

tbm – *Time buffer* da montagem

tbp – *Time buffer* da produção

TOC – *Theory of Constraints*

TR – Tempo de resposta

WIP – *Work-In-Process*

WLC – *Workload Control*

SUMÁRIO

CAPÍTULO 1 - INTRODUÇÃO 16
1.1 Contextualização 16
1.2 Formulação do problema 18
1.3 Objetivos 22
1.4 Relevância do estudo 22
1.5 Estrutura da tese 24
CAPÍTULO 2 – PRODUÇÃO MAKE-TO-ORDER (MTO) 26
2.1 Estratégias de resposta à demanda 26
2.2 Planejamento e Controle da Produção MTO 28
2.2.1 Abordagens clássicas 31
2.2.2 Abordagens emergentes 32
CAPÍTULO 3 - DRUM-BUFFER-ROPE (DBR) 34
3.1 Teoria das Restrições: origem e evolução 34
3.2 *Drum-Buffer-Rope*: conceitos básicos 36
3.3 *Master Production Scheduling* e programação da restrição 38
3.4 Dimensionamento do *time buffer* 41
3.5 Programação da liberação dos materiais 42
3.6 Controle do fluxo da produção: *Buffer Management* 43
CAPÍTULO 4 – CONSTANT WORK-IN-PROCESS (CONWIP) 46
4.1 CONWIP: conceitos básicos 46
4.2 Dimensionamento dos cartões 49
4.3 Sequenciamento dos *jobs* 52
4.4 CONWIP x DBR 53
CAPÍTULO 5 – MÉTODO DE PESQUISA 58
5.1 Caracterização da pesquisa 58
5.2 Etapas da pesquisa 59
5.3 Caracterização do ambiente de produção 61
CAPÍTULO 6 – MODELAGEM CONWIP 65
6.1 Modelo conceitual CONWIP 65
6.2 Modelo computacional CONWIP 69
6.3 Verificação e validação CONWIP 72
CAPÍTULO 7 – MODELAGEM DBR 74
7.1 Modelo conceitual DBR 74
7.2 Modelo computacional DBR 80
7.3 Verificação e validação DBR 82

CAPÍTULO 8 – DELINEAMENTO DOS EXPERIMENTOS 84
8.1 Fatores e variáveis de resposta 84
8.2 Determinação da duração das simulações, tempo de aquecimento e quantidade de replicações 86
8.3 Determinação dos níveis de maxWIP para o modelo CONWIP 90
8.4 Determinação dos níveis de *time buffer* para o modelo DBR 94
CAPÍTULO 9 – RESULTADOS E DISCUSSÕES 98
9.1 Execução da simulação 98
9.2 Resultados do CONWIP 98
9.3 Resultados do DBR 104
9.4 Análise de Variância 109
9.5 Comparação dos resultados dos modelos: CONWIP x DBR 111
9.6 Discussões 116
9.6.1 Questão de pesquisa 1 116
9.6.2 Questão de pesquisa 2 120
CAPÍTULO 10 – CONSIDERAÇÕES FINAIS 122
10.1 Conclusões e contribuições da pesquisa 122
10.2 Limitações do trabalho e desdobramentos 125
REFERÊNCIAS 126
APÊNDICE A – BLOCOS DO *ANYLOGIC* EMPREGADOS NA SIMULAÇÃO 134
APÊNDICE B – ANÁLISE DOS REGIMES TRANSITÓRIO E PERMANENTE DAS VARIÁVEIS 136

CAPÍTULO 1 - INTRODUÇÃO

Este capítulo, inicialmente, contextualiza a temática em que se insere o estudo e, em seguida, formula o problema de pesquisa. Delimitado o problema, são apresentados os objetivos (geral e específicos), a relevância e a estrutura da tese.

1.1 Contextualização

A estratégia de produção contra pedido (*make-to-order* – MTO) ganhou destaque nas últimas décadas em virtude do aumento da demanda por produtos customizados. Muitas empresas migraram da produção repetitiva (*make-to-stock* - MTS) para a MTO ou ainda passaram a adotar uma produção híbrida (STEVENSON; HENDRY; KINGSMAN, 2005; BEEMSTERBOER; LAND; TEUNTER, 2016).

A estratégia MTO é caracterizada por alta variedade e baixo volume dos lotes de produção, orientada pela carteira de pedidos firmados, dependendo menos de previsões da demanda. Isso acarreta menores custos de estoque e maior flexibilidade, contudo, implica em aumento do tempo de resposta ao cliente, menor utilização da capacidade produtiva e maior variedade dos fluxos no chão de fábrica. Consequentemente, traz maior complexidade aos processos de planejamento e controle em relação a produção repetitiva (STEVENSON; HENDRY; KINGSMAN, 2005; GARMDARE; LOTFI; HONARVAR, 2018).

Na literatura de Gestão de Operações, destacam-se algumas abordagens para auxiliar as atividades de programação e controle da produção contra pedido, como: *Constant Work-In-Process* (CONWIP), *Drum-Buffer-Rope* (DBR), *Workload Control* (WLC) e *Paired-cell Overlapping Loops of Card with Autorisation* (POLCA), revisadas por Stevenson; Hendry e Kingsman (2005). As primeiras duas estão no escopo do presente estudo, em virtude das recentes discussões na literatura sobre a aplicação desses modelos na produção MTO.

O CONWIP é um mecanismo de controle da produção, proposto como alternativa ao Kanban para a produção intermitente. É conhecido pela simplicidade, pois efetua apenas o controle da entrada e saída dos lotes em uma linha (completa ou parcial) com base em um limite de estoque em processo. Esse controle do fluxo é feito com sinais de liberação ou cartões comuns a todos os produtos fabricados na linha (PRAKASH; CHIN, 2015; JAEGLER et al., 2017).

Em revisão recente da literatura, Jaegler et al. (2017) afirmam que, após o estudo de Stevenson; Hendry e Kingsman (2005), o CONWIP tem sido discutido com mais frequência como uma alternativa para o controle da produção MTO. Além disso, Jaegler et al. (2017) destacam vantagens do modelo, como: (a) estrutura simples, mecanismo fácil de ser compreendido e aplicado; (b) bom desempenho comparado com outros sistemas e, em alguns casos, superior em relação a vazão, tempo de fluxo, atendimento dos prazos de entrega e níveis de estoque em processo; (c) desempenho superior ao Kanban e ao POLCA na produção MTO.

O *Drum-Buffer-Rope*, mecanismo de programação e controle da produção da Teoria das Restrições, também tem tido destaque na literatura para ambientes com alta variedade e é conhecido como uma poderosa abordagem para controle de sistemas com gargalos[1] (MABIN; BALDERSTONE, 2003; GOLMOHAMMADI, 2015; THÜRER et al., 2017c).

O DBR é composto por três elementos: (a) o tambor (*drum*) representa a restrição, elemento que limita a capacidade do sistema atingir níveis mais altos de retorno em relação aos objetivos da empresa; (b) o pulmão (*buffer*) consiste em uma proteção contra variabilidades; (c) a corda (*rope*) simboliza o controle da liberação dos materiais e das ordens com base na programação. Em virtude da abordagem com foco no gargalo, o DBR também é considerado um mecanismo de simples compreensão e implantação (SCHRAGENHEIM; RONEN, 1990).

O CONWIP e o DBR já foram comparados na literatura, sobretudo, na configuração *flow shop*. Nesse ambiente, há muitos trabalhos que apontam o DBR como superior, há outros que mostram equivalência entre os modelos e, em menor proporção, casos onde o CONWIP foi melhor. Estudos comparativos em ambientes mais complexos, como sistemas com montagem e *job shops*, ainda são escassos na literatura. Dessa forma, aponta-se na literatura a necessidade de investigar diante de quais ambientes e configurações é favorável à aplicação de cada um dos modelos a fim de orientar decisões dos gestores (FRAMINAN; GONZÁLEZ; RUIZ-USANO, 2003; THÜRER et al., 2017b).

Frente a esse contexto, a presente tese realiza uma análise comparativa dos modelos de controle CONWIP e DBR no cenário *make-to-order* em um ambiente com montagem. O tópico a seguir detalhará a formulação do problema de pesquisa.

---

[1]A literatura de Teoria das Restrições conceitua e diferencia gargalo, *Capacity Constraint Resource* (CCR – Recurso com Capacidade Restrita) e restrições, que podem ser qualquer fator que limita o desempenho de um sistema ou organização, como explica Sirikrai e Yenradee (2006). Contudo, neste trabalho, os termos gargalo, restrição e recurso crítico/restritivo foram utilizados indistintamente, exceto quando explicitado.

1.2 Formulação do problema

As diferenças entre os modelos CONWIP e DBR estão sendo discutidas, recentemente, diante do fenômeno da flutuação do gargalo. A flutuação pode ser conceituada como a mudança da estação gargalo na linha (sistema) com o tempo (LAWRENCE; BUSS, 1994). Gerenciar um ambiente com flutuação é considerado um desafio muitas vezes inevitável diante da variabilidade dos sistemas. Com mudanças de *mix* de produção e mesmo em decorrência de melhorias do processo, a utilização dos recursos muda e um novo recurso torna-se gargalo (CRAIGHEAD; PATTERSON; FREDENDALL, 2001; BETTERTON; SILVER, 2012).

No entanto, Thürer e Stevenson (2018) contestam que a avaliação do impacto da flutuação no desempenho de sistemas produtivos com controle da liberação de ordens é negligenciada na literatura. Como essa avaliação é complexa, alguns trabalhos abordaram o problema por meio da análise do impacto da posição do gargalo (THÜRER et al., 2017b; THÜRER; STEVENSON, 2018). Os estudos demonstraram que a análise da posição do gargalo pode ajudar a explicar o efeito da flutuação no desempenho de ambientes com controle da liberação.

Thürer e Stevenson (2018) simularam cenários com a identificação incorreta da posição do gargalo em um *flow shop* puro controlado com o DBR e um modelo de WLC. Como os dois mecanismos eram orientados pelo gargalo, o erro na identificação do recurso crítico impactou negativamente a eficiência de ambos.

Já Thürer et al. (2017b) compararam o desempenho do DBR, CONWIP, WLC e um sistema sem controle de liberação, com a análise de diferentes posições do gargalo em um *flow shop* puro. Apesar do DBR e o CONWIP apresentarem desempenho semelhante em alguns indicadores, ao analisarem detalhadamente os resultados, observou-se que o CONWIP manteve o desempenho independentemente da posição do gargalo. No caso do WLC e do DBR, identificou-se uma perda de desempenho quando o gargalo estava localizado em posições mais ao final da linha. Os autores afirmaram que há, então, evidências de que a posição do gargalo impacta no desempenho desses modelos.

Apesar da análise detalhada do impacto da posição do gargalo, os autores citados consideram apenas estruturas simples de produtos, uma sequência linear de operações que não exigia montagens para finalização. Outros artigos mais antigos já estudaram a posição do gargalo com simulação em ambientes mais complexos e com montagens, como Fry et al. (1987) e Bahouth e Foote (1994), contudo, sem controle da liberação de ordens.

Elftman (1999) foi o único trabalho identificado na literatura que considerou um sistema com montagem e controle do fluxo da produção. O autor analisou o impacto da localização de

duas restrições em um sistema testado com diversas formas de controle, entre elas o puxado-empurrado, que aplicava o CONWIP até o gargalo e nas demais estações adotava o fluxo empurrado.

Thürer et al. (2017b) destacam como possível desdobramento da pesquisa deles a investigação do impacto da posição do gargalo em ambientes com fluxos convergentes ou divergentes, como operações de montagens e desmontagens, subordinados ao controle da liberação. Esse direcionamento também é dado na literatura dos sistemas DBR e CONWIP, discutir como tais modelos podem ser aplicados em ambientes com complexidade maior, como: sistemas com montagem, *job shops* e ambientes com alta variedade de produtos e/ou de roteiros de produção (BETTERTON; COX, 2009; GOLMOHAMMADI, 2015; PRAKASH; CHIN, 2015; JAEGLER et al., 2017).

Os sistemas com montagens são ambientes de produção formados por linhas de fabricação e operações de montagem, que realizam a união dos componentes/módulos oriundos dessas linhas. Esses sistemas podem possuir duas configurações de acordo com as características da fabricação: (a) *assembly shop* (ou *assembly job shop*), uma variação do ambiente *job shop*, onde há diversos roteiros no processo de fabricação e montagem dos componentes (PEREIRA; SANTORO, 2011); (b) *assembly flow shop*, uma variação do *flow shop*, onde a fabricação dos componentes acontece em linhas independentes e o fluxo produtivo tem uma mesma direção (CHEN et al., 2015).

Em ambientes com operações de montagem, observa-se o desafio da sincronização das operações. A sincronização visa a disponibilidade das peças da lista de materiais do produto, fabricadas internamente ou adquiridas de fornecedores, necessárias para a operação de montagem. A sincronização é importante para manutenção do fluxo de produção, atendimento dos prazos de entrega e melhor utilização da capacidade (UMBLE; SRIKANTH, 1990; PEREIRA; SANTORO, 2011).

Apesar de diversos estudos discutirem a programação e aplicação de regras de despacho nesses ambientes com montagens, pouco tem sido discutido sobre o controle da liberação das ordens (LU; HUANG; YANG, 2011). A coordenação das operações de fabricação dos componentes torna-se ainda mais crítica quando se aumenta a carga no sistema, o que torna relevante a discussão sobre como planejar e realizar a liberação de modo a minimizar o tempo de fluxo das ordens (BERTRAND; VAN DE WAKKER, 2002).

Diante do exposto, a presente tese busca estender as discussões de trabalhos anteriores sobre a comparação dos modelos CONWIP e DBR, considerando um ambiente com montagem

(*assembly flow shop*) e diferentes posições do gargalo. A presente pesquisa está direcionada pelas seguintes questões:

Qual é o impacto da posição do gargalo no desempenho do controle efetuado pelo CONWIP e DBR em um ambiente de montagem MTO?

Qual modelo é mais apropriado para o controle de um assembly flow shop que produz contra pedido?

A primeira questão de pesquisa busca analisar se a posição do gargalo (localização do recurso crítico no *layout*) influencia o desempenho desses mecanismos em relação aos indicadores: tempo médio de fluxo, *lead time* médio, percentual de atraso, atraso médio e utilização do gargalo, que serão detalhados no Capítulo 8. Já a segunda questão remete ao desempenho relativo dos modelos, a fim de evidenciar diferenças entre os mesmos no ambiente em estudo.

A consideração de um ambiente com montagem diante do contexto de alta variedade traz alguns desafios para a configuração do CONWIP e DBR (GOLMOHAMMADI, 2015; JAEGLER et al., 2017). Spearman; Woodruff e Hopp (1990) afirmam que o CONWIP pode ser aplicado em sistemas com montagem e, devido à necessidade de sincronização é importante o bom sequenciamento das ordens. Já o DBR foi concebido para a programação e o controle de ambientes com gargalo e montagens, de forma que pressupõe a proteção dessas operações em sua configuração. No entanto, pouco é discutido na literatura acadêmica do DBR sobre avaliações de desempenho em ambientes com montagem (GUIDE JR.; GHISELLI, 1995; YE; HAN, 2008).

O Quadro 1 evidencia alguns indicativos de lacuna na literatura e os pressupostos extraídos de estudos anteriores em consonância com o problema de pesquisa. Cada construto está relacionado a pelo menos um objetivo específico (detalhados no tópico seguinte).

Quadro 1 - Definição do problema de pesquisa

Construtos	Oportunidades de pesquisa	Pressupostos	Objetivos
Posição do gargalo	Não tem sido explorado o impacto da posição do gargalo em ambientes mais complexos com controle de liberação de ordens, incluindo casos com fluxos convergentes e com fluxos divergentes (THÜRER et al., 2017b).	Quando o gargalo é na operação de montagem, há maior comprometimento da vazão do sistema, mas torna-se mais fácil o controle. Quando o gargalo é em uma operação da linha de fabricação, há maior complexidade para o sequenciamento e o controle com o CONWIP (SPEARMAN; WOODRUFF; HOPP, 1990; DUENYAS, 1994a). Quando o gargalo está posicionado no final do processo de fabricação (anterior a montagens), há um maior comprometimento do desempenho. Quando o gargalo está posicionado no início do processo de fabricação, o sistema apresenta melhor desempenho (FRY et al., 1987).	B
CONWIP x DBR	Pesquisas futuras devem comparar o desempenho dos modelos em ambientes com fluxos convergentes e divergentes a fim de analisar diante de quais contextos e fatores a aplicação de cada modelos é mais indicada (THÜRER; STEVENSON, 2018). Em comparações do CONWIP com o DBR, deve-se analisar se o número de máquinas/operações após o gargalo (posição do gargalo), a diferença entre a capacidade de gargalo e a do restante as máquinas (severidade do gargalo) (...) afetam o desempenho desses modelos (FRAMINAN; GONZÁLEZ; RUIZ-USANO, 2003).	Em alguns estudos sobre *flow shop*, o DBR superou o CONWIP (GILLAND, 2002; KIM; DAVIS; COX III, 2003; NG; BERNEDIXEN; SYBERFELDT, 2012). Em Jodlbauer e Huber (2008), o CONWIP apresentou melhores resultados e nas simulações de Thürer et al. (2017b), os modelos se mostraram estatisticamente semelhantes. Na comparação dos modelos em um *flow shop*, o CONWIP manteve o desempenho independentemente da posição do gargalo analisada. Já o DBR apresentou uma perda de desempenho quando o gargalo estava localizado em posições mais ao final da linha. A severidade do gargalo não evidenciou diferenças entre os modelos CONWIP e DBR (THÜRER et al., 2017b).	A, C; D

Fonte: Elaboração própria.

1.3 Objetivos

A presente tese pretende comparar o desempenho dos modelos de programação e controle CONWIP e DBR em um *assembly flow shop* com produção MTO. Em consonância com o objetivo geral desta pesquisa, delineia-se os seguintes objetivos específicos:

a) Investigar como os modelos CONWIP e DBR desempenham o controle do sistema de montagem MTO e quais as diferenças entre eles, diante de indicadores relacionados ao cumprimento de prazo, velocidade de fluxo e utilização da capacidade;
b) Avaliar o efeito da posição do gargalo no desempenho do *assembly flow shop* com controle da liberação de ordens;
c) Analisar como a calibração dos parâmetros de controle (*time buffer* no DBR e limite de WIP no CONWIP) e a severidade do gargalo afetam o desempenho dos modelos no *assembly flow shop*;
d) Identificar qual o mecanismo mais indicado diante dos fatores do experimento.

1.4 Relevância do estudo

Para evidenciar a relevância das questões de pesquisa, realizou-se uma busca na base de dados *Scopus* com a seguinte combinação de termos para os campos título, resumo e palavras-chaves: *"conwip" OR "constant work-in-process" OR "constant work in process" AND "dbr" OR "drum-buffer-rope" OR "drum buffer rope".* Observou-se um total de 18 trabalhos publicados entre 1999 e 16 de dezembro de 2018, sendo um artigo de congresso. Ao realizar uma leitura crítica desses estudos, identificou-se apenas 11 artigos que discutiam os dois métodos, sendo 8 estudos comparativos, 2 trabalhos de proposição de novos modelos e um estudo classificado como exploratório. O Quadro 2 resume as principais características dos trabalhos analisados.

A maior parte dos artigos utiliza a simulação como método de avaliação do desempenho dos modelos. Quanto às características dos experimentos, o *flow shop* é o principal ambiente considerado nos estudos, em virtude da representatividade na manufatura e por possuir menor complexidade para análise. Alguns desses trabalhos consideram nos experimentos diferentes posições do gargalo (KIM; DAVIS; COX III, 2003; NG; BERNEDIXEN; SYBERFELDT, 2012; ROCHA et al., 2015; LEE; SEO, 2016; THÜRER et al., 2017b). Dessa forma, a presente

pesquisa irá contribuir com a comparação do CONWIP e DBR em um ambiente ainda pouco explorado na literatura.

Quadro 2 - Estudos sobre os modelos CONWIP e DBR publicados até 16 de dezembro de 2018

Estudo	Tipo	Objetivo	Modelos	Técnica	Ambiente
Gilland (2002)	C	Analisar o impacto de políticas de liberação no desempenho de células de manufatura e como a escolha dos parâmetros afeta o desempenho das políticas	CONWIP e DBR	Simulação e Teoria das Filas	FS
Kim; Davis e Cox III (2003)	C	Comparar do desempenho do CONWIP, DBR e *Dynamic Flow Control* (DFC) em uma linha genérica	CONWIP, DBR e DFC	Simulação	FS
Koh e Bulfin (2004)	C	Comparar o DBR e o CONWIP em uma linha com base em uma Cadeia de Markov em tempo contínuo	CONWIP e DBR	Teoria das Filas e Otimização	FS
Jodlbauer e Huber (2008)	C	Investigar o desempenho do MRP, Kanban, CONWIP e DBR em um sistema com produtos com vários níveis diante de condições estocásticas	CONWIP, DBR, MRP e Kanban	Simulação e Otimização	FS
Kim; Cox e Mabin (2010)	E	Estabelecer uma estrutura para entender o papel do WIP e da capacidade protetiva em um sistema produção reentrante com *Evaporating Clouds* e simulação	CONWIP, DBR e Kanban	Simulação	FS
Fredendall; Ojha e Patterson (2010)	C	Classificar e comparar regras da teoria de *Workload Control* (WLC), incluindo o CONWIP e DBR	CONWIP, DBR e WLC	Simulação	JS
Ng; Bernedixen e Syberfeldt (2012)	C	Propor uma metodologia para comparação de mecanismos de controle da produção diante de configurações ótimas gerada com fronteiras ótimas de Pareto	CONWIP, DBR, Kanban e um sistema empurrado	Simulação e Otimização multi-objetivo	FS
Millstein e Martinich (2014)	P	Propor um mecanismo para controle de sistemas desbalanceados com alto tempo de ciclo, *setups* e flutuações do gargalo e compará-lo com o CONWIP, DBR e o *one-piece flow*.	CONWIP, DBR, *one-piece flow* e *Takt Time Group*	Simulação	FS
Rocha et al. (2015)	P	Propor um modelo de controle para o fluxo de um sistema híbrido de produção (MTS e MTO)	CONWIP + DBR e *Base-stock* + DBR	Simulação	FS
Lee e Seo (2016)	C	Comparar os modelos DBR, CONWIP e Kanban em um sistema com tempos constantes com a aplicação de métodos exatos	CONWIP, DBR e Kanban	Álgebra Max-Plus	FS
Thürer et al. (2017b)	C	Investigar o impacto da posição do gargalo no desempenho de diferentes modelos de controle de WIP	CONWIP, DBR e WLC	Simulação	FS

Legenda: C – Estudos comparativos; E – Estudos exploratórios; P – Estudos com proposição de novos modelos; FS- *Flow shop*; JS- *Job shop*.
Fonte: Elaboração própria.

A maior parte das pesquisas do Quadro 2 não especifica a estratégia de produção considerada. Apenas Rocha et al. (2015) discutem alternativas para o controle de um sistema de produção híbrido (*make-to-stock* e *make-to-order*). Jodlbauer e Huber (2008) destacam a simulação de um ambiente com vários produtos finais e a variação da demanda desses no experimento, o que pode indicar um sistema de alta variedade. Logo, a presente pesquisa irá contribuir com análise comparativa dos modelos na perspectiva da produção MTO.

Os resultados dos estudos do Quadro 2 ressaltam a importância de se considerar o efeito de diversos fatores para a escolha dos mecanismos de programação e controle. Nas análises comparativas, observou-se que o DBR superou o CONWIP na maioria dos estudos (GILLAND, 2002; KIM; DAVIS; COX III, 2003; KOH; BULFIN, 2004; NG; BERNEDIXEN; SYBERFELDT, 2012).

No entanto, discute-se que a aplicação do DBR depende da correta identificação do gargalo. A cada mudança da posição do gargalo, altera-se a programação e o controle da liberação. Já o CONWIP possui a vantagem em relação ao DBR de não ser tão afetado com a flutuação. Essas diferenças oportunizam a investigação dos efeitos das condições do ambiente no sistema, a fim de identificar se determinados cenários são favoráveis para o bom desempenho e aplicação do CONWIP ou DBR (GILLAND, 2002; FRAMINAN; GONZÁLEZ; RUIZ-USANO, 2003; THÜRER et al., 2017b).

Além da relevância acadêmica da pesquisa, a temática abordada é pertinente para a prática gerencial. Gargalos limitam o rendimento dos sistemas e identificá-los é o primeiro passo para um processo de melhoria. Como discutido na formulação do problema, há indícios de que a posição do gargalo influencia o desempenho dos sistemas de programação e controle da produção. Desse modo, é importante ter a habilidade gerencial de identificar rapidamente qualquer mudança na posição para adotar o controle adequado diante desse novo cenário (THÜRER et al., 2017b).

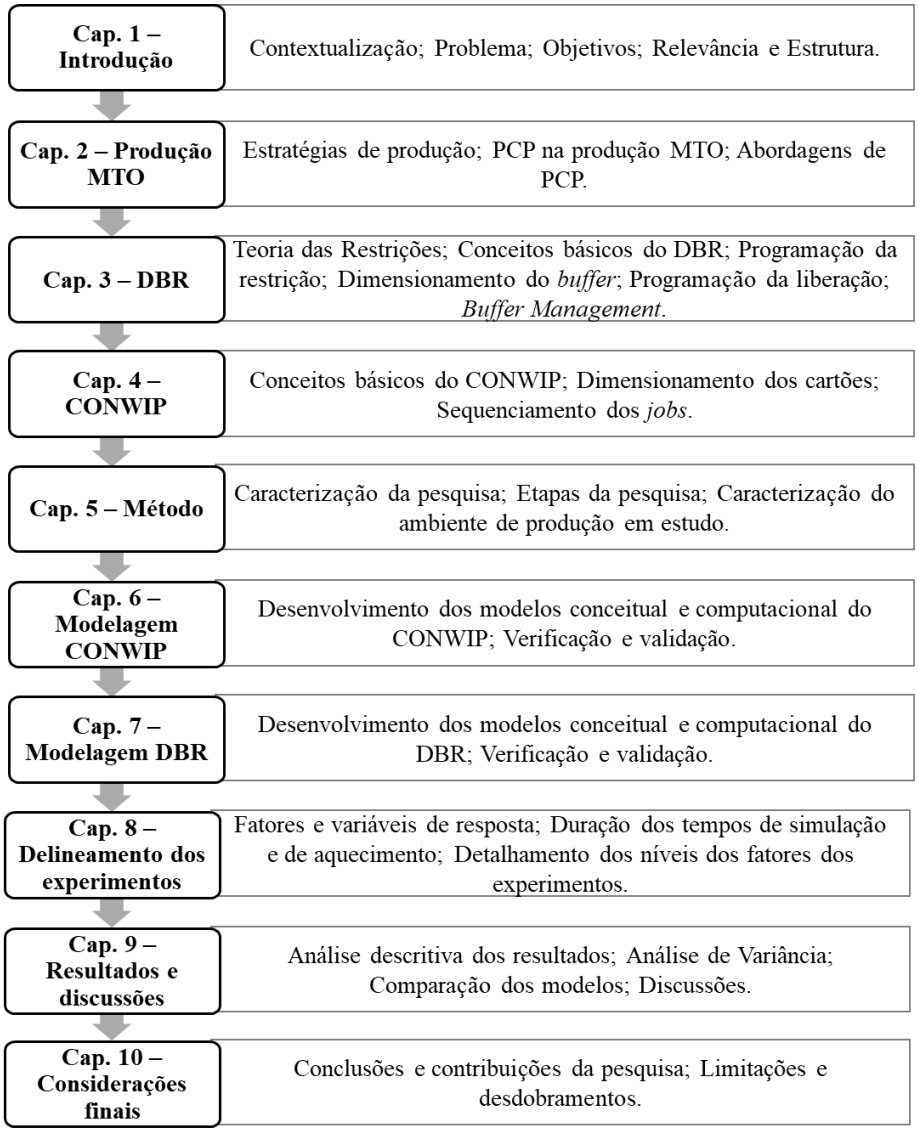
1.5 Estrutura da tese

A tese está organizada em dez capítulos como representado na Figura 1. O Capítulo 1 (Introdução) apresenta a problemática, os objetivos e a relevância do estudo, como detalhado nos tópicos anteriores. O Capítulo 2 (Produção *make-to-order*) explora a estratégia MTO, detalha a estrutura do Planejamento e Controle com foco no processo de gerenciamento de pedidos e faz uma breve revisão dos principais modelos de controle da produção.

Os Capítulos 3 e 4 tratam da revisão da literatura dos modelos *Drum-Buffer-Rope* e *Constant Work-In-Process*, respectivamente. O Capítulo 5 (Método de pesquisa) classifica a abordagem de pesquisa adotada, detalha as etapas para elaboração da pesquisa e descreve as características do ambiente em estudo.

Os Capítulos 6 e 7 apresentam a modelagem conceitual e computacional dos sistemas CONWIP e DBR, respectivamente. O Capítulo 8 detalha os fatores e variáveis de resposta do delineamento dos experimentos, duração das simulações e tempo de aquecimento e os níveis dos fatores. No Capítulo 9 são apresentados os resultados do estudo e feitas as discussões, conforme a problemática da pesquisa. Por fim, no Capítulo 10 (Considerações finais) realiza-se a síntese das conclusões e contribuições do estudo, assim como as limitações e oportunidades de pesquisas futuras.

Figura 1 - Estrutura da tese



Fonte: Elaboração própria.

CAPÍTULO 2 – PRODUÇÃO MAKE-TO-ORDER (MTO)

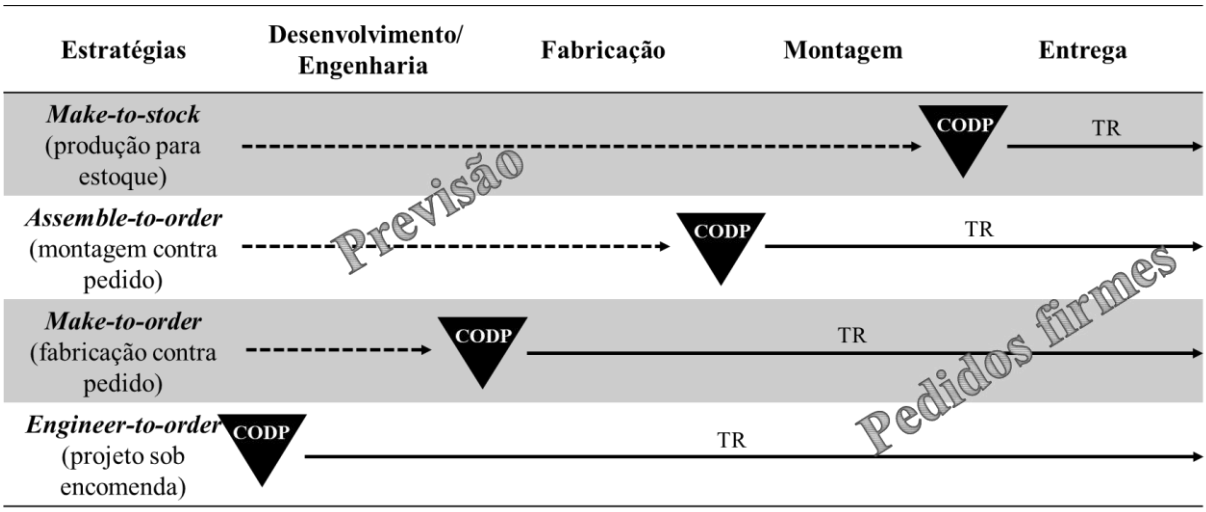
Este capítulo objetiva caracterizar a produção *make-to-order*. Inicialmente, são apresentadas as principais estratégias de resposta à demanda. Em seguida, descreve-se o Planejamento e Controle da Produção no contexto MTO, relacionando-o com processo de gerenciamento dos pedidos. Por fim, as principais abordagens de Planejamento e Controle nesse ambiente são revisadas.

2.1 Estratégias de resposta à demanda

Os sistemas de manufatura podem ser classificados quanto à estratégia de resposta à demanda em quatro tipos principais: *make-to-stock*, *assemble-to-order*, *make-to-order* e *engineer-to-order*. As diferentes estratégias estão relacionadas com o ponto de desacoplamento (*customer order decoupling point* – CODP) (OLHAGER, 2003).

O CODP define em que estágio da cadeia ou do processo, as atividades deixam de ser orientadas por previsões e passam a trabalhar com "pedidos firmes", ou seja, associa-se um determinado produto em fabricação ou acabado a um pedido/ordem de um cliente. Nesse estágio, há um acúmulo de estoque, que determina o tempo de resposta ao cliente (OLHAGER, 2010). A Figura 2 apresenta as principais estratégias e a posição do CODP na cadeia.

Figura 2 - Estratégias de resposta à demanda



Legenda: TR – Tempo de resposta ao cliente.
Fonte: Adaptado de Olhager (2003).

Na estratégia *make-to-stock* (produção para estoque - MTS), os produtos são fabricados com base na previsão da demanda e mantidos em estoque para atender as necessidades dos clientes. Trabalha-se com um alto volume de produção e baixa variedade (produção repetitiva). Diante das incertezas, dimensionam-se estoques de segurança e as decisões são orientadas pelos custos. A previsão da demanda, o gerenciamento dos estoques e o dimensionamento dos lotes são essenciais para essa estratégia (BEEMSTERBOER; LAND; TEUNTER, 2016; ATAN et al., 2017).

Na produção contra pedido (*make-to-order*), foco desta pesquisa, o processo de fabricação é iniciado apenas após o recebimento de um pedido. Oferece maior variedade de produtos e um volume de produção menor, geralmente, exige maquinário/força de trabalho flexível. Em virtude da customização, *lead times* são maiores que a estratégia MTS e a *assemble-to-order*. A promessa de entrega é feita ao cliente com base na capacidade disponível de fabricação (WEMMERLÖV, 1984; STEVENSON; HENDRY; KINGSMAN, 2005).

A montagem contra pedido (*assemble-to-order* - ATO) é uma estratégia intermediária entre MTS e MTO, que busca oferecer customização a baixo custo. Componentes, peças e submontagens padronizadas são produzidas para estoque de acordo com a previsão da demanda. A operação final de montagem é executada apenas quando há um pedido, conforme as especificações do cliente. A promessa de entrega e o planejamento são feitos com base na disponibilidade dos componentes (WEMMERLÖV, 1984; ATAN et al., 2017).

Por fim, em projetos sob encomenda (*engineer-to-order* - ETO), produtos altamente customizados são fabricados em baixa quantidade, de acordo com as especificações do cliente. A cada pedido inicia-se o processo de desenvolvimento do produto e, em seguida, todas as etapas de produção (GRABENSTETTER; USHER, 2017). O Quadro 3 resume as principais características e desafios das estratégias apresentadas.

Quadro 3 - Características das estratégias de resposta à demanda

Estratégia	Customização	Velocidade de entrega	Custo da customização	Complexidade gerencial	Principal desafio
ETO	Produtos únicos (customização total)	Mais lenta	Mais alto	Alta	Projeto e fabricação
MTO	Grande customização	Lenta	Alto	Moderada	Programação (*scheduling*)
ATO	Leve (módulos de montagem)	Rápida	Leve	Leve	Planejamento do estoque de componentes
MTS	Nenhuma	Imediata	Nenhum	Mais baixa	Previsão da demanda

Fonte: Adaptado de Akinc e Meredith (2015).

Na decisão por uma estratégia de produção, observa-se o dilema entre rapidez e variedade. Um sistema que possui apenas estoques de insumos, tem como vantagem uma maior flexibilidade, oferecendo ao cliente produtos customizados, enquanto que um sistema que possui estoques de produtos acabados, ganha como diferencial um atendimento mais rápido (HOPP; SPEARMAN, 2013).

Como pode-se observar no Quadro 3, o aumento da customização adiciona complexidade gerencial, o que implica em desafios diferentes para cada estratégia. Segundo Akinc e Meredith (2015), o principal desafio para um ambiente que opera com a estratégia MTO é a programação. Nesse contexto, Hendry e Kingsman (1989) defendem que a dificuldade de programação pode ser amenizada com um enfoque maior no controle da carga no sistema. Stevenson; Hendry e Kingsman (2005) afirmam que ao controlar a liberação das ordens, pode-se adotar métodos mais simples para a programação. Essas atividades serão discutidas nos tópicos seguintes, tanto na perspectiva do DBR (Capítulo 3), quanto do CONWIP (Capítulo 4).

2.2 Planejamento e Controle da Produção MTO

A decisão por uma determinada estratégia de resposta à demanda afeta a forma como são organizados os processos de fabricação e o Planejamento e Controle da Produção (PCP) (WEMMERLÖV, 1984; BEEMSTERBOER; LAND; TEUNTER, 2016). A literatura ressalta que a estratégia MTO possui particularidades que aumentam a complexidade gerencial em relação a MTS, tais como: incertezas em relação aos roteiros de produção, variações dos tempos de processamento dos pedidos, flutuações da demanda no curto e médio prazo, além de fatores operacionais (KINGSMAN; TATSIOPOULOS; HENDRY, 1989).

A estrutura do PCP no contexto MTO está relacionada com as etapas do gerenciamento de pedidos. Essa estrutura é discutida com maior frequência na literatura de *Workload Control*. Ebadian et al. (2009) afirmam que o planejamento e controle da produção MTO compreende três estágios: negociação e recebimento do pedido (*order entry*), controle da liberação da ordem (*order release*) e sequenciamento/priorização dos *jobs* (*sequencing*/*dispatching*). Nesta tese, no entanto, entende-se que a liberação é antecedida pela etapa de programação das ordens da fila de entrada (*scheduling*), havendo, então, quatro etapas: *order entry*, *order scheduling*, *order release* e *order sequencing*.

No primeiro nível, *order entry* (negociação)¸ é avaliado se o sistema é capaz de receber novas ordens e negociada as condições e o prazo de entrega (SOEPENBERG; LAND; GAALMAN, 2008; EBADIAN et al., 2009). Alguns autores sugerem a segmentação em dois

estágios para entrada da ordem: "*customer enquiry*" e "*job entry*". O primeiro compreende o recebimento da requisição de cotação, a aceitação/rejeição da ordem e a definição da data de entrega e preço do serviço/produto. O segundo envolve a preparação para produção das ordens confirmadas. Após aceitas, as ordens irão compor o *backlog list* ou *pre-shop pool* (STEVENSON; HENDRY; KINGSMAN, 2005; THÜRER; STEVENSON; SILVA, 2011).

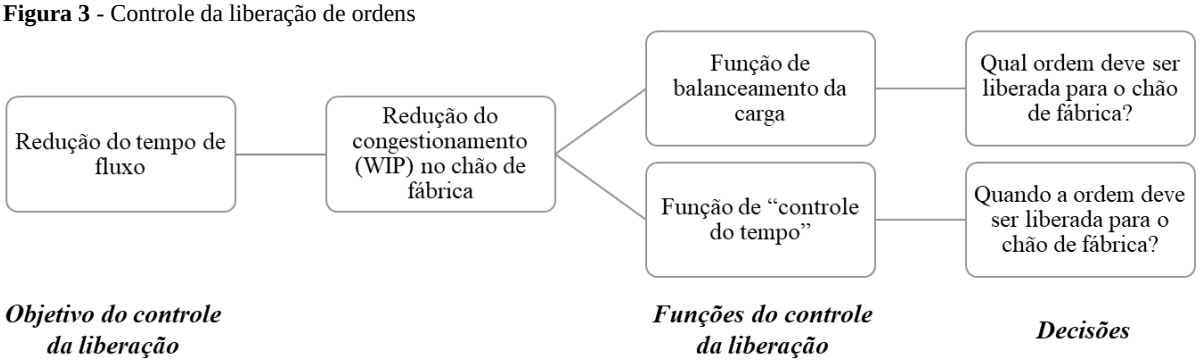
A etapa programação dos *jobs* (*scheduling*) considera uma carteira de pedidos já definida, o *status* da fábrica e a disponibilidade de matérias-primas. A programação visa o mínimo estoque em processo, menores tempos de atravessamento, maior confiabilidade de entrega e melhor utilização da capacidade. Logo, essa atividade é orientada por indicadores que podem ser classificados em três classes: (a) cumprimento de prazos; (b) velocidade de fluxo; (c) utilização da capacidade do gargalo (HOPP; SPEARMAN, 2013).

A programação da produção é a atividade do PCP que tem sido mais discutida na literatura nos últimos anos. Entre as soluções discutidas estão algoritmos de otimização, heurísticas e regras de prioridade. Muitas vezes em virtude da dificuldade de soluções ótimas e da dinâmica da demanda, as regras acabam sendo a opção mais frequente na prática.

Alguns autores defendem a aplicação de regras simples para o sequenciamento das ordens, acompanhadas de um controle de carga mais restritivo para a liberação (HENDRY; KINGSMAN, 1989; EBADIAN et al., 2009; THÜRER; STEVENSON; SILVA, 2011). Em mecanismos, como o *Simplified Drum-Buffer-Rope*, sugere-se que ao invés da programação detalhada dos recursos, haja o enfoque no controle do fluxo da produção, conforme necessidade da demanda (SOUZA; BAPTISTA, 2010).

O controle da liberação das ordens objetiva a redução do tempo de fluxo, por meio da limitação do WIP, o que proporciona a melhoria do controle da qualidade do processo com a redução do congestionamento no chão de fábrica. A decisão da liberação, normalmente, está condicionada a algum fator, seja segundo a carga de trabalho dos recursos (relacionada a capacidade) ou em função do momento planejado/programado para liberação, de acordo com o prazo de entrega (BERGAMASCHI et al., 1997; EBADIAN et al., 2009). A Figura 3 relaciona o objetivo do controle da liberação com as decisões feitas no estágio de *order release*.

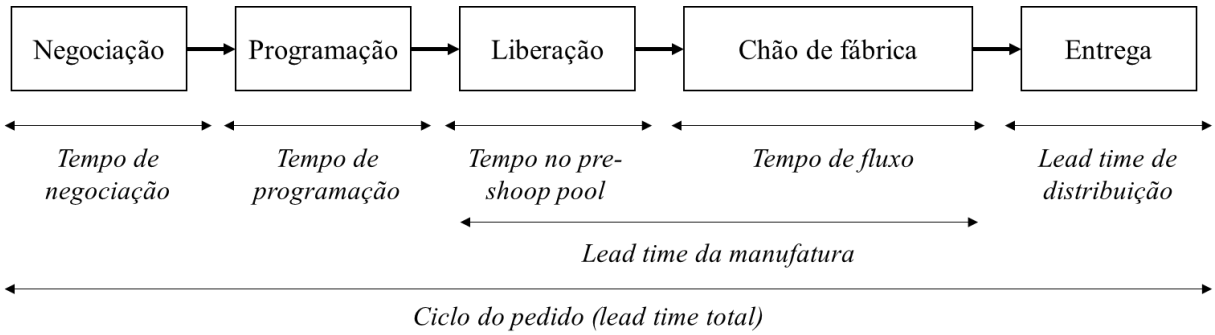
Figura 3 - Controle da liberação de ordens



Fonte: Adaptado de Ebadian et al. (2009).

O último nível refere-se ao sequenciamento das ordens nas filas das máquinas no chão de fábrica com o objetivo de reduzir o tempo na fila, de modo que não haja comprometimento do prazo de entrega. A Figura 4 apresenta as principais fases do planejamento e controle da produção no cenário MTO, relacionando-as com o processo de gerenciamento de pedidos e os tempos que compõem o ciclo do pedido.

Figura 4 – Processo de gerenciamento de pedidos



Fonte: Adaptado de Ebadian et al. (2009).

O bom desempenho de entrega é o fator crítico de sucesso da estratégia MTO. Diante da complexidade do ambiente produtivo da fabricação sob encomenda, a escolha do sistema de controle da produção é importante para auxiliar tanto a estimativa dos prazos, como o gerenciamento do *lead time* ao longo das demais etapas do ciclo do pedido (STEVENSON; HENDRY; KINGSMAN, 2005; EBADIAN et al., 2009).

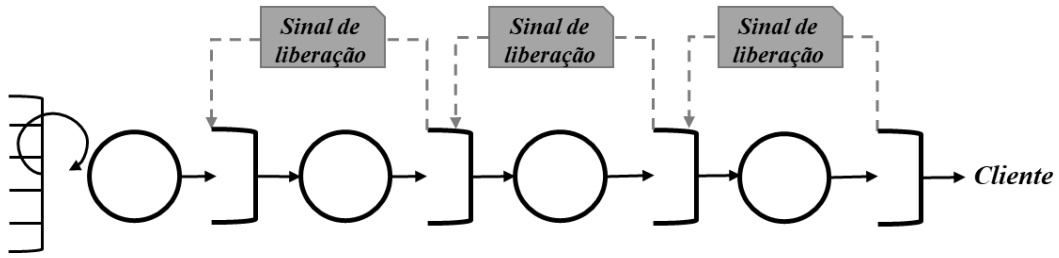
Stevenson; Hendry e Kingsman (2005) revisaram e discutiram a aplicabilidade de abordagens consideradas clássicas (Kanban, *Material Requirements Planning* - MRP e DBR) e "emergentes" (*Workload Control, Paired-cell Overlapping Loops of Cards with Authorization* – POLCA e CONWIP) para o planejamento e controle da produção contra pedido. Esses sistemas serão brevemente discutidos nas próximas seções.

2.2.1 Abordagens clássicas

O Kanban, mecanismo de controle pertencente ao Sistema Toyota de Produção, é a principal referência para o controle puxado da produção. O Kanban é inspirado na reposição de prateleiras de supermercados: um processo cliente retira um produto conforme a sua necessidade, essa retirada autoriza o processo fornecedor a produzir a mesma quantidade e tipos dos itens consumidos. A fabricação e a movimentação dos produtos são autorizadas apenas com cartões ou sinais de liberação específicos para as peças (OHNO, 1988).

A Figura 5 exemplifica o controle executado pelo Kanban. Cada círculo representa uma operação. A quantidade de estoque em processo entre as operações corresponde a quantidade cartões Kanban entre as estações. Ao ser retirado um produto do estoque da última operação, o cartão com as informações do item consumido dispara o processo de reposição. Essa liberação desencadeia a produção das estações anteriores até o fornecimento de insumos.

Figura 5 - Representação do Kanban



Fonte: Elaboração própria.

Essa abordagem foi desenvolvida com foco na produção repetitiva, apesar de amplamente conhecida, não é considerada apropriada para lidar com a variedade de produtos e roteiros característicos da produção MTO (STEVENSON; HENDRY; KINGSMAN, 2005). Como destacam Lage Junior e Godinho Filho (2010), diante de algumas condições a adoção da versão original do Kanban não é indicada: ambientes com tempos de processamento ou demanda instáveis, operações não padronizadas, *setups* longos, variedade de produtos e suprimento não confiável.

O *Material Requirements Planning* (MRP I) é um sistema desenvolvido para auxiliar a programação da produção, definindo o que produzir, quanto e quando produzir, em ambientes com alta complexidade. O MRP I contribuiu com a simplificação da gestão, contudo, os planos desenvolvidos algumas vezes apresentavam lacunas em relação as necessidades de capacidade produtiva. Com isso, esse sistema evoluiu para *Manufacturing Resource Planning* (MRP II) e, posteriormente, para o *Enterprise Resource Planning* (ERP), que integra diversas áreas

funcionais, incluindo o módulo da produção (STEVENSON; HENDRY; KINGSMAN, 2005; FERNANDES; GODINHO FILHO, 2011).

Stevenson; Hendry e Kingsman (2005) destacam a ampla adoção dos sucessores do MRP em diversas empresas no mundo, assim como, relatos de dificuldades para implantação e parametrização dos sistemas. Os autores afirmam que os sistemas não consideram adequadamente as decisões de capacidade na negociação (*order entry*) e no controle da liberação (*order release*), mas que podem ser utilizados na produção MTO em conjunto com mecanismos de controle da produção com cartões como CONWIP, POLCA, etc.

Além do Kanban, MRP I e as evoluções decorrentes desse sistema, o DBR também é classificado como uma abordagem clássica de controle da produção (detalhada no Capítulo 3). Em relação a aplicabilidade do DBR em ambientes com produção contra pedido, Stevenson; Hendry e Kingsman (2005) criticam a dificuldade do mecanismo lidar com a flutuação do gargalo e a ausência de suporte as atividades do nível de negociação das ordens. Apesar disso, destacam a simplicidade do controle efetuado e ressaltam que é uma abordagem amplamente conhecida por gestores, adotada em ambientes de produção altamente customizados.

2.2.2 Abordagens emergentes

Em relação as abordagens chamadas de “emergentes”, Thürer; Stevenson e Silva (2011) explicam que o termo *Workload Control* (WLC) refere-se a um conjunto de métodos decorrentes da união de quatro correntes de pesquisa, que a partir da década de 90 começaram a ser citados como WLC: (a) métodos de revisão e liberação de ordens (*order review and release methods*); (b) controle da carga de trabalho baseado no controle *input*/*output*; (c) controle de fabricação orientado pela carga (*load oriented manufacturing control*); (d) regras para liberação de pedidos.

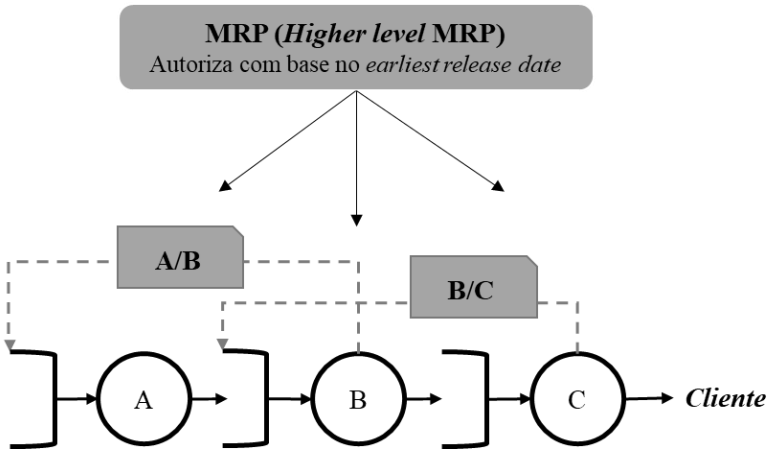
Thürer; Stevenson e Silva (2011) afirmam que os modelos de WLC regulam a entrada no chão de fábrica conforme a carga de trabalho atual, os limites de carga e as características dos *jobs*. A abordagem foca na produção contra pedido e possui modelos que compreendem os diversos níveis do processo de gerenciamento dos pedidos: *job entry*, *order release* e *dispatching*. As principais críticas quanto ao WLC referem-se a aplicabilidade dos conceitos, que são bastante discutidos em simulações e experimentos na literatura, mas há poucos estudos sobre implantações (HENDRY; HUANG; STEVENSON, 2013).

O POLCA é um sistema criado como alternativa ao Kanban no contexto da estratégia *Quick Response Manufacturing* (QRM). É considerada uma abordagem híbrida (puxada e

empurrada), que engloba o controle por cartões e o sistema MRP com foco em ambientes com pequenos lotes e alta variedade. Os cartões controlam a produção e movimentação entre pares de estações/células, enquanto o sistema MRP determina as datas planejadas para produção das ordens com base na estrutura simplificada dos produtos e nos *lead times* das células (CHINET; GODINHO FILHO, 2014).

Para processamento de uma ordem em uma célula "A" sob controle do POLCA é preciso: (a) disponibilidade da célula "A" para o processamento; (b) um cartão referente as células "A/B" (atual e a próxima do roteiro de produção), que indica disponibilidade de capacidade futura na célula "B"; (c) ter atingido a data planejada para produção da ordem na estação "A". No controle com POLCA, os cartões permanecem com as ordens desde a fila da primeira célula até término do processamento na segunda célula e não são específicos de um produto (THÜRER; STEVENSON; PROTZMAN, 2016). A Figura 6 representa a lógica do controle com POLCA em três células/estações (A-B-C).

Figura 6 - Representação do sistema POLCA



Fonte: Adaptado de Thürer; Stevenson e Protzman (2016).

Na literatura, observam-se críticas quanto a aplicabilidade do POLCA em ambientes sujeitos a variedade de roteiros de produção e ao gerenciamento dos cartões. Em algumas situações pode haver um bloqueio do fluxo das ordens, especialmente, na produção MTO (HARROD; KANET, 2013; THÜRER; STEVENSON; PROTZMAN, 2016).

Em relação a adoção do CONWIP na produção MTO, Stevenson; Hendry e Kingsman (2005) afirmam que é preciso certa padronização dos produtos para permitir o controle simplificado de WIP e criticam o uso do modelo original em ambientes com muita variação de fluxo, como *job shops*. Apesar da avaliação dos autores, alguns estudos mais recentes mostram o CONWIP como uma boa opção para essa estratégia e ambiente (HARROD; KANET, 2013; JAEGLER et al., 2017).

CAPÍTULO 3 - DRUM-BUFFER-ROPE (DBR)

Este capítulo apresenta os principais conceitos relativos ao *Drum-Buffer-Rope*. Inicialmente, é feito um breve relato sobre como a Teoria das Restrições (*Theory of Constraints* – TOC) surgiu e evoluiu ao longo dos anos. Em seguida, são apresentados os principais elementos que compõem o DBR e discutidos os passos para aplicação. O *Buffer Management* (BM) também é apresentado como elemento importante para a implantação do DBR.

3.1 Teoria das Restrições: origem e evolução

A origem da Teoria das Restrições está associada ao desenvolvimento do *software Optimized Production Timetables* criado por Eliyahu M. Goldratt em 1979 e, posteriormente, renomeado para *Optimized Production Technology* (OPT). O OPT era um *software* de programação finita, que possuía módulos para o planejamento hierárquico da produção, responsáveis pela elaboração do Programa Mestre de Produção, programação detalhada da restrição e a liberação das ordens (FRY; COX; BLACKSTONE JR., 1992; WATSON; BLACKSTONE; GARDINER, 2007).

Diversas empresas adotaram o *software* e conseguiram bons resultados, mas enfrentaram muitos problemas com a implantação. Por essa razão, Goldratt investiu na divulgação dos nove princípios do OPT e na escrita de livros sobre a filosofia de gestão proposta (SPENCER; COX, 1995; WATSON; BLACKSTONE; GARDINER, 2007).

O livro "*The Goal*" (A meta), publicado em 1984, descreve fundamentos da TOC, como os cinco passos de focalização (*Five Focusing Steps*), que orientam o processo de implantação da teoria. Os cinco passos são:

a) Identificar a restrição do sistema: as restrições determinam o desempenho da empresa, por isso é preciso identificá-las para dar início ao processo de melhoria (WATSON; BLACKSTONE; GARDINER, 2007). Segundo Schragenheim e Ronen (1990), a restrição pode ser um recurso, o mercado consumidor, o fornecimento de algum insumo ou políticas gerenciais da empresa;
b) Explorar ao máximo a restrição: no caso das restrições gerenciais, deve-se eliminá-las e substituí-las por políticas que auxiliem o aumento do ganho (*throughput*). Para as restrições físicas, deve-se buscar a melhoria dos processos para manter o fluxo de

produção na restrição por meio de esforços como o controle da qualidade, rotinas de manutenção preventiva, redução dos *setups* no recurso crítico e uma boa programação. Explorar também significa que cada minuto da restrição deve ser usado na fabricação dos produtos que maximizam o ganho da empresa (RAHMAN, 1998; COX III; SPENCER, 2002);

c) Subordinar todas as demais atividades à restrição: mesmo que os recursos não críticos possuam capacidade além da restrição, esses devem estar alinhados com a programação e a necessidade da restrição para não gerar estoques desnecessários (RAHMAN, 1998);
d) Elevar a restrição do sistema: após explorar a capacidade da restrição (passo b), deve-se empregar alternativas para aumentá-la, como adquirir mais um recurso, modificar o equipamento existente, etc. (RAHMAN, 1998; COX III; SPENCER, 2002);
e) Se nas etapas anteriores a restrição for quebrada, voltar ao primeiro passo: caso a restrição mude, deve-se iniciar novamente o ciclo em busca da melhoria contínua.

Em decorrência dos cinco passos de focalização e das metáforas adotadas em "A meta", surgiu o método de programação *Drum-Buffer-Rope* (em português, Tambor-Pulmão-Corda), apresentado e explicado na publicação seguinte "*The Race*" (1986). O DBR operacionaliza os três primeiros passos de focalização no chão de fábrica e abrange apenas restrições de mercado e recursos produtivos (foco desta pesquisa). Mais recentemente, foi desenvolvido o método *Simplified Drum-Buffer-Rope* (S-DBR), que assume o mercado como a única restrição do sistema (LEE et al., 2010; SOUZA; BAPTISTA, 2010). Esse modelo ainda não é amplamente discutido na literatura acadêmica, mas representa uma alternativa diante das dificuldades de implantação do DBR.

Outra ferramenta proposta por Goldratt para o gerenciamento operacional foi o *Buffer Management* (BM), que consiste em um mecanismo de monitoramento do DBR para auxiliar na tomada de decisão no curto e médio prazo (SCHRAGENHEIM; RONEN, 1991; SPENCER; COX, 1995).

Em 1987, Goldratt começou a adotar o nome "Teoria das Restrições" para se referir a sua proposta de gerenciamento das organizações. Ao longo dos anos, a TOC evoluiu de um *software* de programação da produção para um conjunto integrado de ferramentas que abrange produção, métodos de solução de problemas, gestão de projetos, gestão da cadeia de suprimentos, etc. (RAHMAN, 1998; SOUZA, 2005; IKEZIRI et al., 2018).

Muitos estudos sobre a TOC são apresentados na literatura acadêmica e não acadêmica, destacando os benefícios dessa teoria (MABIN; BALDERSTONE, 2003). Ikeziri et al. (2018)

é a revisão da literatura mais recente sobre essa temática, que detalha a evolução da TOC tanto na perspectiva histórica, quanto por meio de uma análise bibliométrica da produção científica. A revisão demonstra que mesmo após quase quatro décadas desde os primeiros estudos, a temática continua relevante e com oportunidades de pesquisa.

3.2 *Drum-Buffer-Rope*: conceitos básicos

O *Drum-Buffer-Rope* é o mecanismo de programação e controle da produção da Teoria das Restrições. O primeiro elemento do DBR, *drum* (tambor), representa a restrição, que dita o ritmo de produção do sistema. O DBR propõe uma programação focada no gargalo. Os recursos não-restritivos anteriores ao gargalo são subordinados à programação da restrição por meio controle da liberação de ordens, enquanto que os localizados após o gargalo trabalham de forma "empurrada". Essa simplificação é possível, pois os recursos não-restritivos possuem capacidade superior ao gargalo, denominada capacidade protetiva (*protective capacity*). A capacidade produtiva da empresa é, então, determinada pela disponibilidade da restrição (BLACKSTONE JR; COX III, 2002).

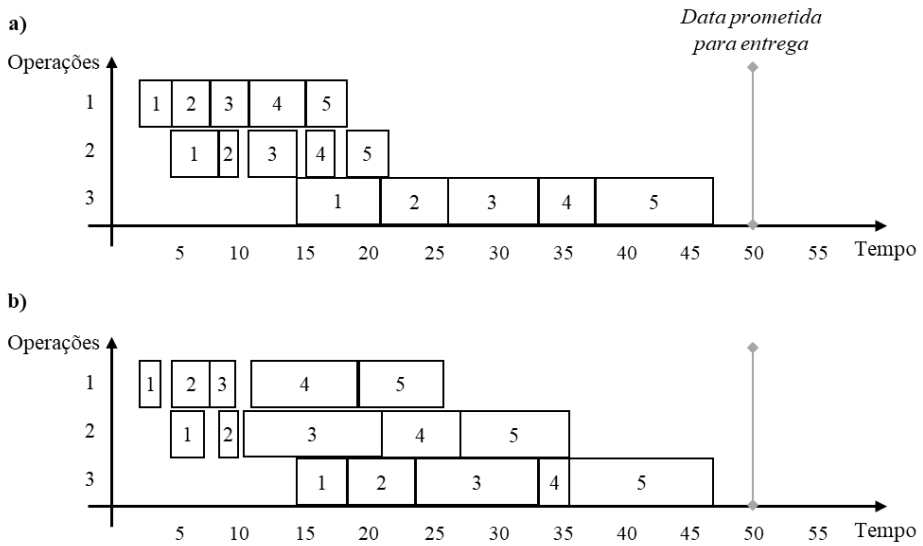
O segundo elemento, *buffer*, garante o fluxo de *jobs* no processo e o atendimento da demanda. O pulmão, denominado na literatura clássica de DBR como *time buffer*, é uma proteção de tempo contra variabilidades, a fim de que os materiais estejam disponíveis no recurso gargalo antes do momento necessário para o processamento (GOLDRATT, 1990; SCHRAGENHEIM; RONEN, 1990). Segundo Goldratt (1990), é mais adequado adotar a proteção em função de tempo do que em unidades de produto, pois o foco é o atendimento das ordens no prazo e isso torna o conceito aplicável a outros contextos, como gestão de projetos.

A Figura 7 exemplifica a aplicação do *time buffer* (tb). Ao programar o gargalo, incorpora-se uma folga para garantir que mesmo diante das incertezas os materiais/*jobs* estejam disponíveis no momento esperado para o processamento no recurso crítico. Na Figura 7a, observa-se a programação de cinco ordens em três máquinas (M1, M2 e M3), sendo a última um gargalo. Nesse exemplo, as ordens foram programadas para chegar com antecedência no recurso crítico e a programação do gargalo foi feita de modo a finalizar todos os pedidos antes da data de entrega.

Na execução da programação (Figura 7b), há variação dos tempos de processamento, o que acarreta a finalização de todas as ordens após o momento estimado na programação

(determinística) para as máquinas 1 e 2. No exemplo, o *time buffer* apesar de ter sido "consumido" com as variações, protegeu a utilização do gargalo e a entrega na data prometida.

Figura 7 - Proteção do *time buffer* à programação do gargalo



Fonte: Elaboração própria.

O DBR em sua concepção inicial considera até três tipos de *buffer*: expedição (remessa), montagem e, o anteriormente exemplificado na Figura 7, para proteger a restrição (*buffer* do gargalo) (RONEN; STARR, 1990). O pulmão de expedição busca garantir o atendimento do prazo de entrega. Segundo Goldratt (1990), o *buffer* da expedição representa o adiantamento da conclusão das ordens e do envio do pedido (se possível). O *buffer* de montagem evita que peças provenientes do gargalo aguardem por peças fabricadas em recursos não-restritivos. Logo, um estoque das peças oriundas de recursos não-restritivos é posicionado antes de operações de montagem (SCHRAGENHEIM; RONEN, 1991; SOUZA, 2005).

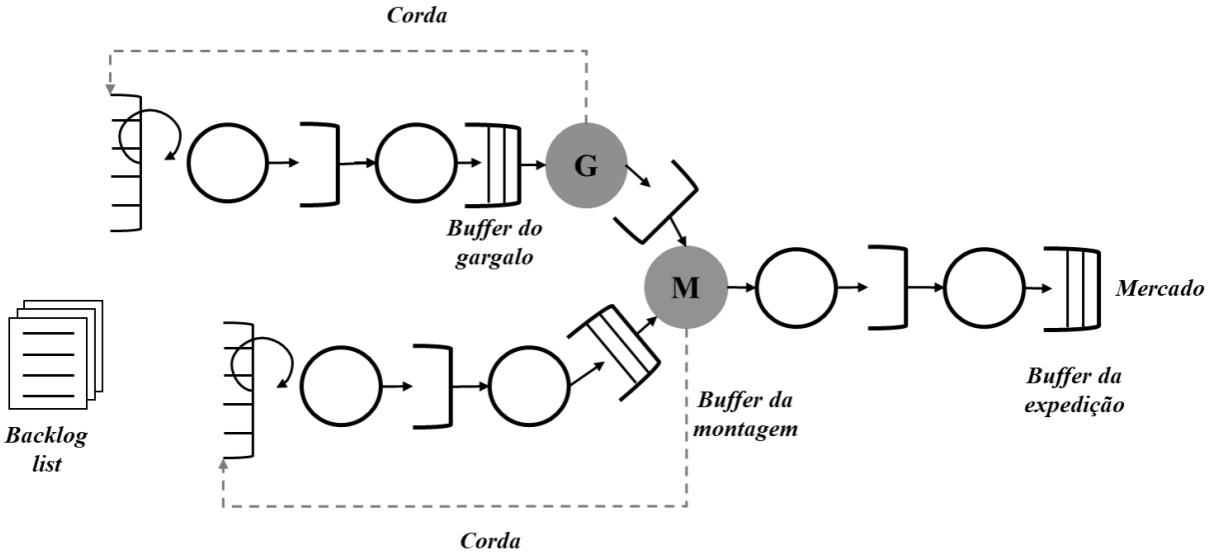
O terceiro elemento, a corda (*rope*), é o elo de *feedback* ao processo, que promove a subordinação das operações à restrição. A corda realiza o controle da liberação dos materiais para que os recursos não-restritivos não produzam mais do que o programado. Busca-se manter os níveis de WIP mínimos e constantes conforme o consumo da(s) restrição(ões) (SCHRAGENHEIM; RONEN, 1990; SOUZA; BAPTISTA, 2010).

Segundo Cox III e Schleier Jr (2010), a corda deve ser empregada em um pequeno número de pontos de controle. Em cada ponto, deve haver uma programação detalhada com informações sobre os materiais a serem liberados, a sequência e o momento planejado para tal

ação. De acordo com o processo e o tipo de restrição (física ou mercado), pode haver configurações distintas do DBR e, consequentemente, diferentes pontos para monitoramento.

Quando o mercado é a restrição, a corda controla a liberação de matéria-prima e entrada dos *jobs* programados, conforme o consumo do *buffer* da expedição. Quando há recursos restritivos na produção, aplica-se o controle segundo o *buffer* do gargalo. Em ambientes com operações de montagem que sucedem a restrição, a corda também restringe a entrada de materiais na linha "livre" de acordo com o consumo do *buffer* da operação de montagem. A Figura 8 exemplifica os três *buffers* do DBR em uma linha com uma operação gargalo e montagem.

Figura 8 - Estrutura básica do DBR com recurso restritivo e montagem



Fonte: Adaptado de Watson; Blackstone; Gardiner (2007).

Schragenheim e Ronen (1990) resumem as etapas para a aplicação do DBR em três passos: (a) programar a restrição, o que implica em desenvolver simultaneamente o Programa Mestre de Produção (*Master Production Scheduling* - MPS); (b) dimensionamento dos *time buffers*; (c) gerar a programação da liberação dos materiais com base nos passos anteriores. Estes passos são detalhados nas seções a seguir.

3.3 *Master Production Scheduling* e programação da restrição

A elaboração do Programa Mestre da Produção compreende os dois primeiros passos de focalização da TOC (identificar e explorar a restrição) e busca a maximização do lucro da empresa. O desenvolvimento do MPS é um problema conhecido na literatura da TOC como

"*product mix heuristics*". A heurística da TOC apresentada em Goldratt (1990) é resumida em Fredendall e Lea (1997).

As referências clássicas sobre o DBR, em sua maioria, não fazem distinção entre as estratégias de produção. Quando se trabalha com produção *make-to-stock*, os passos discutidos por Fredendall e Lea (1997) se mostram aplicáveis. Contudo, quando a produção é sob encomenda, o MPS é definido pelos pedidos aceitos e a programação é direcionada, por exemplo, pela urgência das ordens (THÜRER; STEVENSON, 2017). O S-DBR, em contrapartida, já possui uma abordagem que diferencia as estratégias de produção (SOUZA, 2005; LEE et al., 2010).

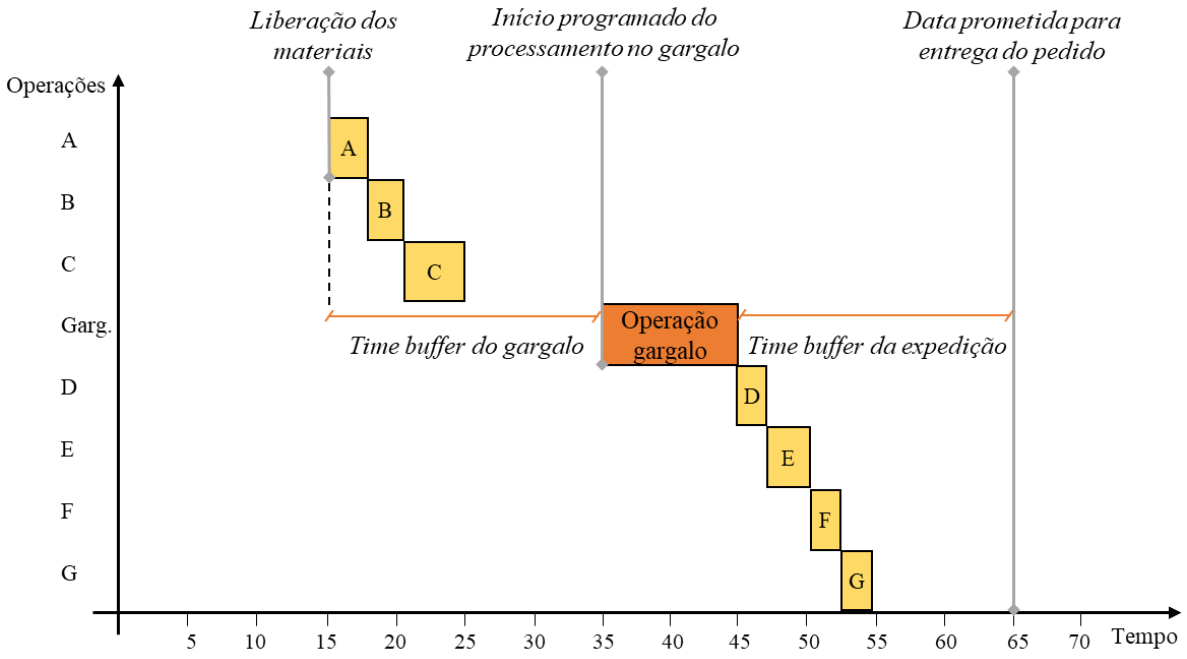
Após a definição do que produzir e quanto produzir por previsão (ambiente MTS), ou com uma carteira de pedidos (ambiente MTO), realiza-se a programação da restrição (*drum*), que definirá quando produzir. No DBR, adota-se o carregamento finito dos recursos com a programação "para trás", considerando os prazos de entrega dos pedidos, e "para frente" quando ocorrem conflitos ou sobreposições na programação, seguindo alguma regra de prioridade (CHAKRAVORTY; ATWATER, 2005; THÜRER; STEVENSON, 2017).

Para determinar a programação da restrição, isto é, os momentos de início do processamento de cada *job* no recurso gargalo, deve-se subtrair do prazo de entrega do pedido, a soma do valor do *buffer* da expedição e o tempo de processamento no gargalo. Todos os recursos não críticos seguem a programação do gargalo. Quando as ordens programadas são liberadas, segue-se normalmente a regra FCFS (*first come, first served*) (COX III; SPENCER, 2002; CHAKRAVORTY; ATWATER, 2005).

A Figura 9 exemplifica a lógica da programação de um processo de fabricação composto por oito operações (A-B-C-Operação Gargalo-D-E-F-G). O produto possui como data prometida para a entrega o momento 65. O *time buffer* da expedição e do gargalo foram definidos como 20 unidades de tempo cada e a operação gargalo dura 10 unidades. Assim, se a entrega deve ocorrer no momento 65, o processamento no gargalo deve iniciar no momento 35 (65-20-10=35). As operações seguintes (D-E-F-G) trabalharão conforme o processamento do gargalo.

Definido o início planejado no recurso crítico, subtrai-se o valor do *time buffer* do gargalo para determinar o momento de liberação dos materiais e início de processamento na linha, ou seja, 35-20=15. Logo, a ordem deve ser liberada no momento 15 para processamento nas operações não críticas que antecedem o gargalo (A-B-C). No exemplo, espera-se que a operação C seja concluída no instante 25 e permaneça no pulmão do gargalo até o processamento.

Figura 9 - Programação do gargalo e da liberação dos materiais



Fonte: Elaboração própria.

Simons Jr. e Simpson III (1997) detalham os passos para programação de múltiplas restrições e explicam o conceito de *rods* (folgas planejadas para garantir o cumprimento da programação das restrições). A lógica descrita por Simons Jr. e Simpson III (1997) se assemelha com a explicação de Souza (2005) para aplicação do DBR no contexto de um sistema informatizado. No entanto, o processo de programação do DBR com múltiplas restrições é considerado de complexa execução (GILLAND, 2002) e não será considerado no escopo deste estudo.

Discussões mais recentes apontam que o desempenho do DBR é bastante afetado pela programação. Thürer e Stevenson (2017) advogam que o procedimento para determinar a programação no DBR pouco mudou ao longo dos anos e recebeu pouca atenção na literatura. Em consonância com os autores, Golmohammadi (2015) ressalta a necessidade de revisão dos procedimentos para programação no DBR, em especial as recomendações para definição dos tamanho dos lotes e do tamanho do *buffer* para aplicação em sistemas reais ou complexos.

Observam-se, portanto, oportunidades de estudos na literatura que contribuam para a programação da restrição. Sobretudo, como enfatizam Thürer et al. (2017c), é importante investigar como a programação pode ser melhorada no DBR sem comprometer a simplicidade característica do método.

3.4 Dimensionamento do *time buffer*

A determinação do tamanho dos *buffers* é uma decisão importante para o desempenho do sistema. Quando subdimensionado, o pulmão pode não cumprir com a função de proteger o gargalo de eventuais interrupções, ou seja, podem deixá-lo esperando por novos trabalhos para processar. Em contrapartida, quando o pulmão é muito grande há aumento nos tempos de atravessamento e acumulação de estoque em processo (RADOVILSKY, 1998; COX III; SCHLEIER JR, 2010).

O dimensionamento do *time buffer* está associado ao nível de flutuações da produção e da capacidade protetiva dos recursos não-restritivos. Quanto mais variabilidade no processo ou quanto menor a diferença entre a capacidade do gargalo em relação aos demais recursos, maior a necessidade de proteção. O tamanho dos *buffers* é um parâmetro do sistema utilizado na programação do gargalo e no controle da liberação dos materiais, por isso, não deve mudar com frequência.

Poucos estudos, contudo, abordam o problema de dimensionamento na literatura e encontrar um valor ótimo para o tamanho do pulmão é reconhecidamente uma tarefa difícil (RADOVILSKY, 1998; YE; HAN, 2008).

Dentre as pesquisas analíticas que abordaram tal problemática, Radovilsky (1998) utilizou conceitos de Teoria das Filas para proposição de um tamanho ótimo do pulmão da restrição. Considerou-se um modelo de filas com chegada Poisson, tempos de atendimento exponenciais, único servidor (recurso restritivo) com fila finita com capacidade K (*M/M/1/K*) e taxa de utilização de 100%. O dimensionamento proposto buscou a maximização do lucro operacional (*net profit*) com base nos valores de lucro unitário, custo unitário para manter os estoques e a taxa de atendimento.

Tu e Li (1998) e Ye e Han (2008) analisaram o problema de dimensionamento do *buffer*, assumindo que a proteção é necessária devido às falhas nas máquinas que precedem a restrição. Os dois estudos sugerem a análise do fluxo dos materiais e das relações de precedências das máquinas como primeiro passo do algoritmo para determinação do *time buffer*. As formulações matemáticas para dimensionamento do pulmão incorporam a taxa de falhas e a função densidade de probabilidade do processo de reparo de cada máquina.

Ye e Han (2008) ampliaram o trabalho de Tu e Li (1998), propondo formulações para determinar também o tamanho do pulmão da montagem de um *assembly shop*. O estudo de Ye e Han (2008) concluiu que quanto maior o número de máquinas precedentes à operação de

montagem, maior o tamanho de *buffer* da montagem. Em contrapartida, quanto maior o número de máquinas que precedem o recurso crítico, menor o tamanho do *buffer* do gargalo.

Louw e Page (2004) desenvolveram um procedimento para estimativa do *time buffer* com base em um modelo de rede de filas aberta com *m* servidores (*GI/G/m*). Os autores estimaram o *time buffer* como a soma do tempo médio de fluxo até o *buffer* e ***z*** (nível de serviço em função do desvio padrão do tempo de fluxo), menos a soma dos tempos de processamento. Para demonstrar a validade do procedimento, os autores compararam as estimativas de *time buffer* com a formulação de filas com os resultados da simulação de um *flow shop* e um *job shop*.

Muitos trabalhos, no entanto, sugerem uma determinação empírica do *time buffer*, a ser posteriormente ajustada conforme o desempenho da produção monitorado com o *Buffer Management*. Schragenheim e Ronen (1990) indicam para ambientes estáveis o valor de três vezes o *lead time* médio até a restrição como tamanho inicial do *buffer* do gargalo. Já Fry; Karwan e Steele (1991) ao discutirem implantações do DBR sugerem que o pulmão deve ser dimensionado como três vezes o desvio padrão do tempo de atravessamento até a restrição.

Blackstone Jr (2013) recomendam que o tamanho do *buffer* seja inicialmente equivalente à metade do *lead time* do sistema, adotando igual valor para o pulmão da restrição e da expedição. Essa é a orientação mais atual para o dimensionamento dos *time buffers*. Segundo Blackstone Jr (2013), há uma relação que deve ser observada entre o dimensionamento dos pulmões (da restrição, montagem e expedição) e o tempo de atravessamento.

Em virtude das diversas orientações existentes na literatura, adotar uma regra empírica (*rule of thumb*) ou deixar essa decisão a cargo da experiência dos gestores pode suscitar insegurança na implantação do DBR. Corbett e Csillag (2001) analisaram empresas brasileiras que adotaram o DBR e identificaram que algumas não implantaram todos os *buffers* como recomendado na literatura, comprometendo o desempenho de entrega. Além disso, uma das principais dificuldades encontradas, relatada em todos os casos analisados, foi o gerenciamento dos pulmões (CORBETT; CSILLAG, 2001).

3.5 Programação da liberação dos materiais

A decisão de liberação das ordens, “corda”, está associada à elaboração da programação da restrição e ao dimensionamento do *buffer*. O momento para liberação dos *jobs* e dos materiais é determinado diminuindo o *time buffer* da programação definida para o gargalo e no caso de linhas com operação de montagem a liberação é dependente do *time buffer* da montagem (SOUZA; BAPTISTA, 2010; GOLMOHAMMADI, 2015).

Ao investigar o elemento corda e as diferentes maneiras implantá-la, Russell e Fry (1997) simularam três estratégias de *order review/release* existentes na literatura clássica de *job shop* a fim de avaliar os impactos no desempenho do DBR. As abordagens testadas foram:

a) Carregamento infinito para trás (*Backward infinite loading*): liberação dos *jobs* de acordo com a programação, um mecanismo de liberação empurrado que não considera as condições do chão de fábrica e de tempo, ou a capacidade dos recursos;
b) Tempo real: liberação dos materiais de acordo com o monitoramento contínuo da taxa de processamento do gargalo (mecanismo puxado). A liberação de um novo lote para a linha ocorre quando uma quantidade equivalente de trabalho for processada no gargalo;
c) Carregamento finito para trás (*Backward finite loading*): é feita uma previsão dos tempos em que os materiais devem ser liberados de acordo com uma programação para trás do gargalo.

Além das políticas de liberação, o estudo de Russell e Fry (1997) analisou o impacto dos lotes de processamento e de transferência serem idênticos ou diferentes e da severidade do gargalo. O controle da liberação em tempo real apresentou melhores resultados na simulação independentemente da política de lotes de transferência/processamento e do balanceamento dos recursos. A abordagem de carregamento finito, característica do DBR, apresentou bom desempenho quanto a entrega, especialmente, quando há diferentes lotes de processamento e transferência.

A proposta de controle em tempo real se mostra semelhante ao CONWIP, sendo que com o controle da carga até gargalo ao invés de toda a linha de produção. Observa-se que alguns estudos sobre DBR adotam o controle da liberação com base no WIP como o CONWIP, ao invés de considerar o *lead time* de segurança (*time buffer*). Esse controle por *jobs* é uma simplificação utilizada em diversos trabalhos de simulação.

3.6 Controle do fluxo da produção: *Buffer Management*

O gerenciamento dos pulmões é considerado pela literatura da TOC um mecanismo importante para aplicação e o bom funcionamento do DBR. A corda é o elemento responsável pelo controle de entrada de novos *jobs* no processo, enquanto que o *Buffer Management* monitora a execução da programação. Ao comparar o plano com a situação atual do chão de

fábrica, o BM indica quais ordens ou lotes precisam de medidas corretivas (SCHRAGENHEIM; RONEN, 1991; SRIKANTH, 2013).

Segundo Schragenheim e Ronen (1991) e Ikeziri et al. (2018), o gerenciamento dos pulmões serve para: (a) alertar possíveis problemas urgentes que podem comprometer a programação; (b) priorizar as ordens; (c) indicar ao planejamento a necessidade de revisão de parâmetros, como o tamanho do pulmão; (d) identificar as principais causas de atraso; (e) indicar áreas do chão de fábrica que precisam de maior atenção e ações de melhoria.

Para identificar a necessidade de ajustes, utiliza-se um indicador de "*status* do pulmão", que avalia quanto cada ordem ou lote de produção consumiu da reserva (pulmão) existente na programação. O indicador relaciona o tempo disponível para a ordem (quanto tempo há entre o momento atual e a data prometida para entrega - TD) e o tempo de atravessamento programado, que seria equivalente ao próprio pulmão (TA) (Equação 3.1). Esse cálculo é feito periodicamente para orientar as ações da gerência (SRIKANTH, 2013).

$$Status\ do\ pulmão\ = \frac{TD}{TA} \times 100 \tag{3.1}$$

Em seguida, as ordens são classificadas em três categorias, conforme criticidade do *status*: verde, amarelo e vermelho. Os limites entre cada região e o valor mínimo, que notifica a necessidade de intervenção, são definidos empiricamente. Srikanth (2013) sugere os seguintes limites e orientações para as três regiões:

a) Região verde: composta pelas ordens que apresentam *status* acima de 67%, ou seja, ainda possuem tempo remanescente superior a dois terços do pulmão e, por isso, não demandam ações corretivas.
b) Região amarela: formada pelas ordens com criticidade intermediária, cujo *status* está entre 33% e 67%. Apesar de não haver ainda um comprometimento do prazo de entrega dessas ordens, existe possibilidade de eventuais variabilidades afetarem o término das operações remanescentes, portanto, deve-se monitorá-las.
c) Região vermelha: constituída pelas ordens com indicador inferior a 33%, ou seja, que possuem tempo remanescente inferior a um terço do pulmão. Essas ordens são críticas, se essas ainda estiverem nos estágios iniciais, deve-se agir corretivamente.

Para assegurar o atendimento do prazo, o processamento das ordens na restrição deve ser priorizado conforme a criticidade (*status*). Além da priorização no sequenciamento, pode-se adotar outras medidas para ajuste do plano como: utilização de horas extras, divisão do lote

para processamento em mais de um recurso ou para o processamento em uma próxima operação à medida que as unidades são concluídas no recurso anterior, etc. (HADAS; CYPLIK; FERTSCH, 2009; SRIKANTH, 2013).

O BM é considerado uma ferramenta de melhoria contínua da produção, uma vez que fornece *feedback* sobre o dimensionamento do *time buffer* e outros parâmetros do processo (IKEZIRI et al., 2018). Se frequentemente houver uma quantidade considerável de ordens na região vermelha (mais de 15%), necessitando de ações corretivas, isso pode indicar que o tempo de atravessamento estimado não está adequado e o *time buffer* deve ser aumentado. O inverso também precisa ser avaliado, se houver muitas ordens na região verde, o pulmão pode sido superestimado e precisa ser gradativamente reduzido (HADAS; CYPLIK; FERTSCH, 2009; SRIKANTH, 2013). Srikanth (2013) sugere que as ordens vermelhas devem representar cerca de 10% das ordens totais.

CAPÍTULO 4 – CONSTANT WORK-IN-PROCESS (CONWIP)

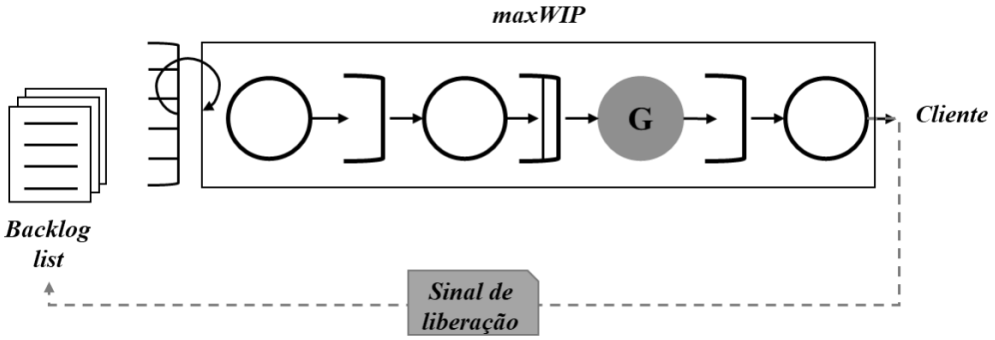
Este capítulo apresenta os principais conceitos relativos ao *Constant Work-In-Process*. Inicialmente, descreve-se o funcionamento do modelo conforme a versão original. Em seguida, são detalhadas as duas principais decisões para configuração e implantação do CONWIP: o dimensionamento de cartões e o sequenciamento do *backlog list*.

4.1 CONWIP: conceitos básicos

O *Constant Work-In-Process* (CONWIP) é um mecanismo de controle da produção, introduzido por Spearman; Woodruff e Hopp (1990) como alternativa ao Kanban, para garantir o fluxo puxado em ambientes de alta variedade. O CONWIP controla o fluxo produtivo por meio de sinais ou cartões, que limitam o estoque em processo no sistema. É considerado, atualmente, o método de controle com cartões mais simples que há na literatura (PRAKASH; CHIN, 2015; THÜRER; STEVENSON; PROTZMAN, 2016).

A Figura 10 apresenta a lógica do CONWIP. Cada *job* em processamento no sistema é sinalizado com um cartão. Ao término do processamento de um *job* na linha, libera-se a entrada de uma nova ordem. Para a liberação deve-se seguir a sequência estabelecida no *backlog list* (relação de trabalhos aguardando a liberação). Com esse controle, o CONWIP mantém um limite máximo de WIP no sistema, denominado neste estudo de "maxWIP". O controle pode ser efetuado para toda linha, como representado na Figura 10, ou para um conjunto de estações. A delimitação do controle é definida pelos *loops*. Após a liberação, as ordens são, normalmente, processadas segundo FCFS (SPEARMAN; WOODRUFF; HOPP, 1990).

Figura 10 - Representação da lógica do CONWIP em uma linha

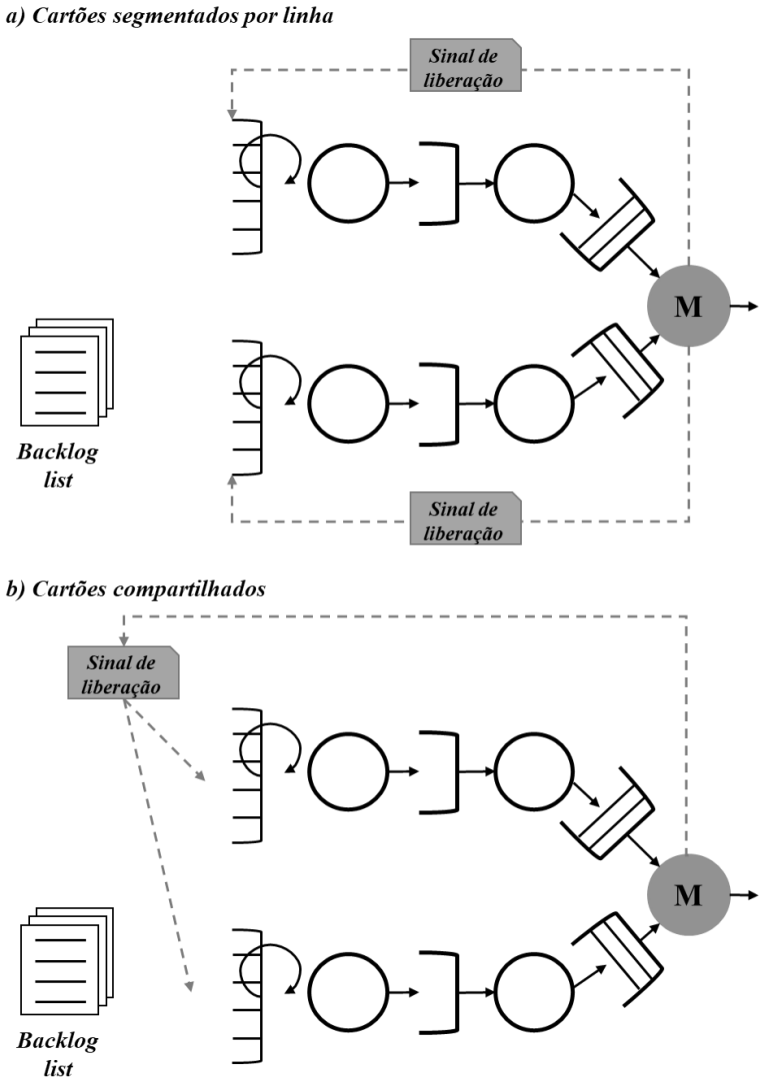


Fonte: Elaboração própria.

O CONWIP foi originalmente desenvolvido para ambientes de produção simples, com roteiros de produção e tempos de processamento semelhantes. Diferentemente do Kanban, os cartões no CONWIP não são específicos para cada produto ou itens em processamento, mas são compartilhados para cada linha ou fluxo. Spearman; Woodruff e Hopp (1990), ao introduzirem o modelo, afirmaram que esse também era aplicável em ambientes com montagem e gargalos.

Na presença de operações de montagem, o CONWIP pode auxiliar a coordenação das linhas de fabricação com o controle do estoque em processo nas linhas (Figura 11). Khojasteh-Ghamari (2009) avaliou duas possibilidades de controle nesse cenário: (a) cartões segmentados por linha e (b) cartões compartilhados. O estudo concluiu que a política de cartões segmentados apresentou resultado superior: menor quantidade de WIP médio e maior taxa de *throughput*.

Figura 11 - CONWIP em linhas com montagem



Fonte: Adaptado de Spearman; Woodruff e Hopp (1990) e Khojasteh-Ghamari (2009).

Spearman; Woodruff e Hopp (1990) discutem ainda o impacto de um recurso gargalo no ambiente com montagem. Os autores afirmam que se a montagem for o gargalo ou se nessa estação houver significativos *setups*, espera-se que o CONWIP tenha bom desempenho. Contudo, se o gargalo estiver nas linhas de fabricação e houver *setups* altos, haverá mais dificuldade para sequenciar adequadamente os *jobs* e garantir uma boa utilização do sistema.

Em ambientes desbalanceados, em especial na configuração *flow shop*, Spearman; Woodruff e Hopp (1990) defendem que o CONWIP se assemelha ao DBR. Como o ritmo do recurso mais lento (gargalo) determina a utilização da linha, os demais recursos eventualmente permanecerão ociosos, aguardando o processamento. O DBR, no entanto, pode ser considerado mais genérico do que o CONWIP (em sua versão original), pois é aplicável em ambientes mais complexos, como *job shops*. A principal diferença entre os modelos, segundo os autores, diz respeito ao momento de liberação, que no CONWIP ocorre automaticamente conforme a taxa do gargalo.

Ao longo dos anos foram criadas diversas variações do modelo CONWIP original para ampliar a sua aplicabilidade. Há na literatura modificações do modelo para considerar cartões para cada família de produto, regras de despacho dinâmicas, controle do maxWIP em estações intermediárias, etc., como detalha a revisão de Prakash e Chin (2015).

O CONWIP é considerado um método robusto, mais flexível e fácil de implantar que outros mecanismos de controle da produção puxada (FRAMINAN; GONZÁLEZ; RUIZ-USANO, 2003). Para parametrização do CONWIP é importante definir:

a) Quota de produção (*production quota* - a meta de produção para um determinado período, equivalente ao *takt time* do Sistema Toyota de Produção);
b) Produção máxima;
c) Indicador de falta de capacidade (*capacity shortage trigger*);
d) Previsão do *backlog list*;
e) Quantidade de cartões para operação do sistema (maxWIP);
f) Sequenciamento dos *jobs*.

As decisões (a), (b), (c) e (d) pertencem ao nível tático do PCP e devem ser tomadas ao desenvolver um planejamento hierárquico da produção, enquanto que as duas últimas são atividades de controle (SPEARMAN; WOODRUFF; HOPP, 1990; FRAMINAN; GONZÁLEZ; RUIZ-USANO, 2003).

A maior parte das pesquisas discutem o dimensionamento dos cartões, principal parâmetro para configuração do CONWIP, e o sequenciamento dos *jobs*. O sequenciamento pode não ser crítico em situações com produção MTS, mas ganha importância em ambientes

de alta variedade (MTO) (FRAMINAN; GONZÁLEZ; RUIZ-USANO, 2003). As discussões acerca dessas decisões são detalhadas nas seções a seguir.

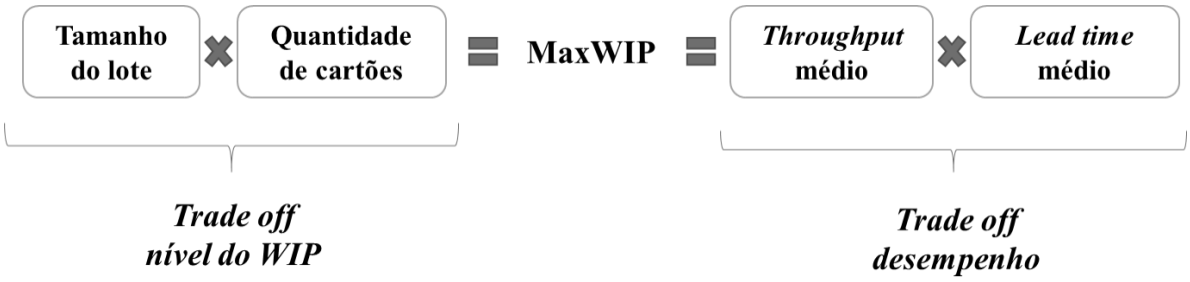
Segundo Spearman; Woodruff e Hopp (1990), os parâmetros para o CONWIP podem ser otimizados, contudo, deve-se observar que há normalmente *trade-offs*. Os autores sugerem adotar parâmetros iniciais conservadores e, posteriormente, reduzi-los a medida em que o desempenho da linha é avaliado.

4.2 Dimensionamento dos cartões

A definição do nível de WIP (maxWIP) está relacionada com a quantidade de cartões que irão circular no sistema e o tamanho dos lotes. Os estudos sobre dimensionamento dos cartões, normalmente, determinam o nível de WIP com base um desempenho desejado em termos de *throughput*, como estabelece a Lei de Little: WIP médio = *throughput* médio x tempo de fluxo ou *lead time* médio, ou outro objetivo para o sistema (NG; BERNEDIXEN; SYBERFELDT, 2012; DURMUSOGLU; AGLAN, 2018).

A relação entre o nível de WIP e o desempenho, assim como os *trade offs* envolvidos no processo de configuração do CONWIP, foram representados por Jaegler et al. (2017) ao revisar a literatura do modelo (Figura 12). Como discutido na literatura, um aumento da quantidade de cartões, eleva a utilização dos recursos e, consequentemente, a produtividade (*throughput*) do sistema. Contudo, isso acarreta um efeito negativo no tempo de fluxo (*lead time*). Para reduzir esse tempo, que impacta na resposta ao cliente, é preciso reduzir o nível de WIP (quantidade de cartões) (KIM; COX; MABIN, 2010; HOPP; SPEARMAN, 2013; JAEGLER et al., 2017).

Figura 12 - Decisão de dimensionamento do WIP



Fonte: Adaptado de Jaegler et al. (2017).

Hopp e Spearman (2013) orientam que para a implantação do CONWIP em uma linha já em funcionamento, a melhor abordagem inicial para o dimensionamento deve ser manter o nível de WIP atual e, posteriormente, reduzi-lo com base na observação das filas e desempenho

da máquina gargalo. Para linhas a serem implantadas ou reconfiguradas, sugere-se a análise com simulação das relações entre o nível de WIP, o tempo de fluxo e a produtividade. O nível de WIP deve ser estabelecido acima da demanda, mas de forma equilibrada.

Para uma linha com operação de montagem (como um sistema com montagem em um estágio representado na Figura 11a), Hopp e Spearman (2013) afirmam que a equação da produtividade nesse sistema (*TH*) pode ser usada para o cálculo dos cartões para cada segmento (w), assumindo a taxa de processamento do gargalo do sistema (g) e o nível crítico de WIP (W_0). Esse nível crítico representa o nível de WIP sob circunstâncias ideais (determinísticas) em cada linha, com base na soma dos tempos de processamento da linha de fabricação (T_0), logo:

$$TH(w) = \frac{w}{W_0 + w - 1} \times g \tag{4.1}$$

$$W_0 = g \times T_0 \tag{4.2}$$

Spearman; Woodruff e Hopp (1990) afirmam que a determinação do nível de WIP no CONWIP é equivalente ao dimensionamento do *time buffer* do DBR. Enquanto o DBR utiliza o *time buffer* para garantir o processamento contínuo do gargalo, no CONWIP, um estoque físico (*inventory buffer*) se acumulará na frente do gargalo devido as diferentes taxas de processamento dos recursos.

Em revisão da literatura, Framinan; González e Ruiz-Usano (2003) observaram que não há regras bem definidas para a determinação da quantidade de cartões. Segundo os autores, há duas decisões principais concernentes ao dimensionamento: *card setting* e *card controlling*. A primeira objetiva determinar a quantidade adequada de cartões para a operação (abordagem estática do problema). A segunda busca o desenvolvimento de regras para orientar o ajuste na quantidade de cartões de acordo com eventos ou variações do sistema (abordagem dinâmica).

O Quadro 4 sistematiza os principais estudos que objetivaram a determinação de uma quantidade ótima de cartões (*card setting*). Além dos estudos citados no Quadro 4, há estudos com otimização e simulação que ao analisarem, compararem ou modificarem a proposta inicial do CONWIP também estipularam uma quantidade de cartões para o sistema, de acordo com a observação do desempenho do modelo em relação a algum objetivo.

Como pode ser observado no Quadro 4, a maior parte dos estudos aplicaram conceitos de Teoria das Filas para determinar o nível de WIP do sistema e o foco está em equilibrar o impacto do WIP como sugere a Lei de Little.

Quadro 4 - Sistematização das abordagens para o dimensionamento dos cartões

Estudo	Ambiente	Objetivo	Técnica
Braglia et al. (2011)	*Flow shop* com gargalo, variação dos lotes de processamento nas máquinas e tempos determinísticos	Minimização do WIP em relação a um determinado nível de *throughput*	Modelo analítico
Duenyas; Hopp e Spearman (1993)	*Flow shop* com consideração de quebras de máquinas e tempos determinísticos	Minimização dos custos	Teoria das Filas
Duri; Frein e Lee (2000)	*Flow shop* com consideração de produtos defeituosos	Minimização de custos	Teoria das Filas
Huang; Wang e Ip (1998)	Linhas com compartilhamento de recursos, gargalo e diversos produtos	Minimização do WIP em relação a um determinado nível de *throughput*	Modelo analítico e simulação
Ioannidis e Kouikoglou (2008)	*Flow shop* com único tipo de produto	Maximização do lucro líquido	Teoria das Filas
Ryan; Baynat e Choobineh (2000)	*Job shop* com múltiplos produtos	Minimização do WIP em relação a um determinado nível de *throughput*	Teoria das Filas

Fonte: Adaptado de Framinan; González e Ruiz-Usano (2003) e Durmusoglu e Aglan (2018).

Outra vertente da literatura de CONWIP, com menor quantidade de publicações, propõe um modelo adaptativo de acordo com a perspectiva de *card controlling*. Prakash e Chin (2015) afirmam que o ajuste dinâmico da quantidade de cartões traz mais proteção para o sistema, além de permitir um melhor desempenho em relação a abordagem estática de gerenciamento (*card setting*).

Hopp e Roof (1998) foram os primeiros a propor um método para ajuste do WIP específico para o CONWIP no contexto de produção *make-to-order*, chamado *Statistical Throughput Control* (STC), baseado no Controle Estatístico do Processo. Com base em um nível de *throughput* alvo, monitora-se o desempenho do sistema diante dos limites de controle e se necessário aumenta-se ou reduz-se a quantidade de cartões.

Framinan; González e Ruiz-Usano (2006) propuseram um sistema de ajuste dos cartões também com base no monitoramento do nível de *throughput* (nível de serviço), sendo verificado a cada *job* concluído no sistema (MTO) ou a cada chegada (no ambiente MTS). Uma quantidade extra máxima de cartões foi estipulada para o ajuste (adição ou redução no maxWIP) a fim de não ocasionar "estresse" no sistema com as variações. O modelo superou o STC em experimentos simulados, os autores ressalvam apenas os impactos que o aumento do WIP e, consequente aumento do tempo de fluxo, apresenta em ambientes MTO no atendimento as datas prometidas.

Posteriormente, Liu e Huang (2009) construíram um procedimento para controle da quantidade de cartões que considera a diferença entre o valor de *throughput* alvo e o valor real para o ajuste, sabendo que o gargalo é um fator limitante da produtividade. O modelo proposto

foi comparado com o de Framinan; González e Ruiz-Usano (2006) e apresentou melhores resultados.

Mais recentemente, Belisário e Pierreval (2015) propuseram uma abordagem com algoritmo genético e simulação para dinamicamente determinar a mudança da quantidade de cartões. Os autores destacam que, diferentemente dos estudos anteriores, essa abordagem gera uma lógica de decisão que especifica diante de quais circunstâncias deve-se modificar a quantidade de cartões (aprendizado automático). Além disso, Jaegler et al. (2017) apresentam uma revisão mais ampla sobre como os estudos lidam com a decisão de dimensionamento (tanto na abordagem estática como dinâmica) dos cartões no CONWIP.

4.3 Sequenciamento dos *jobs*

O sequenciamento dos *jobs* na fila de entrada do sistema é um dos problemas mais discutidos na literatura do CONWIP, conhecido como "*backlog-sequencing problem*" ou "*CONWIP sequencing problem*". O sequenciamento é considerado importante para melhoria do desempenho do sistema, em especial, ao auxiliar no balanceamento de carga em contextos com grande variedade de *mix* (THÜRER et al., 2017a).

Dois fatores importantes para o sequenciamento são destacados por Framinan; González e Ruiz-Usano (2003): (a) estratégia de produção; (b) roteiros com gargalo. Apesar de na prática ser difícil a existência de uma empresa puramente MTO ou MTS, os objetivos dessas estratégias são distintos, o que afeta o processo de sequenciamento. Em relação ao gargalo, deve-se considerar o impacto da sequência nos tempos de *setup* dessa máquina. Caso não haja *setup* dependente, simplificações podem ser feitas, assumindo, por exemplo, o problema de programação de máquina única.

Diversos estudos foram desenvolvidos ao longo dos anos para otimização da sequência em relação a algum objetivo com a aplicação de formulações de Programação Linear (PL)/Não Linear (PNL), heurísticas e meta-heurísticas (FRAMINAN; RUIZ-USANO; LEISTEN, 2001; CAO; CHEN, 2005; AJORLOU; SHAMS, 2013; AGLAN; DURMUSOGLU, 2015). Um dos primeiros estudos publicados sobre a temática na literatura foi Herer e Masin (1997). Os autores formularam um modelo de PL para programação dos *jobs* do *backlog list* de um *flow shop* com o objetivo de minimização dos custos totais, que incluíam custos de manutenção de estoques (produtos acabados e WIP), custo de falta e horas-extras.

Além dos estudos de otimização, outros trabalhos se concentraram na análise, comparação e, em alguns casos, proposição de novas regras de despacho (DUENYAS, 1994b; FRAMINAN; RUIZ-USANO; LEISTEN, 2000; LEU, 2000; BAHAJI; KUHL, 2008; THÜRER et al., 2017a). Nesses trabalhos, discute-se a aplicação de regras de prioridade como uma alternativa para uma programação dinâmica dos *jobs* em cenários que não trabalham com previsão da demanda dos produtos ou quando o *backlog list* está sujeito a constantes alterações.

Segundo Thürer et al. (2017a), apesar de nem todos os estudos com CONWIP explicitarem como o *backlog list* foi sequenciado, as regras mais aplicadas na literatura são FCFS (*first come, first served*), EDD (*Earliest Due Date*) e PRD (*Planned Release Date*). Contudo, em virtude da simplicidade do controle do CONWIP, orientado pela taxa de saída dos *jobs* do sistema, os autores afirmam que o emprego da PRD e outros métodos que programem o momento para liberação podem não ser necessários.

Durmusoglu e Aglan (2018), ao revisarem a literatura sobre CONWIP, afirmam que boa parte dos estudos foca no aprimoramento do mecanismo de liberação dos *jobs*, com o desenvolvimento de novas regras e abordagens híbridas, ao invés do sequenciamento ou programação das ordens. Isso demonstra que a literatura tem buscado uma melhoria do mecanismo de controle para permitir uma simplificação da atividade de programação, em virtude da dinâmica a que está sujeita a demanda.

Há uma dependência entre as atividades de programação e controle, que deve ser observada para auxiliar o bom desempenho da execução da programação. Como comenta Hopp e Spearman (2013), o limite de WIP estabelecido pelo mecanismo CONWIP permite adiantar os *jobs* caso o processamento da linha seja mais rápido do que o esperado, ou seja, se o nível de WIP diminuir, pode-se liberar um *job* antes do momento programado. Como também, caso haja um atraso na programação, o modelo impede a liberação de mais *jobs* para o sistema, consequentemente, a redução das filas e do tempo de fluxo das ordens.

4.4 CONWIP x DBR

Diante da revisão sobre o CONWIP e DBR realizada neste capítulo e no anterior, pode-se fazer uma síntese comparativa dos sistemas.

Quanto ao controle, o DBR é um sistema de controle e programação da produção com foco no gargalo, enquanto que o CONWIP foca no *output* da linha/do sistema produtivo sob controle. Em relação a programação das ordens, no DBR realiza-se a programação do gargalo,

levando em consideração o horizonte de planejamento das ordens. São deduzidos os *time buffers* e os tempos de processamento do recurso crítico das datas de entrega das ordens para determinar os momentos programados para processamento no gargalo. As ordens são priorizadas de acordo com alguma regra de despacho.

No CONWIP, não há exigência de programação das ordens. O mais comum é o sequenciamento do *backlog list* de acordo com alguma regra de prioridade de interesse da empresa. Apesar disso, há na literatura trabalhos que aplicam modelos de otimização para programação da fila de entrada.

Após liberadas para o chão de fábrica, as ordens são priorizadas no DBR de acordo com o indicador *status* do pulmão. Em ambientes com roteiros de produção mais simples como *flow shops* puro, em que as ordens possuem a mesma sequência de operações, costuma-se adotar a regra FCFS como simplificação. Em ambientes com diversos fluxos como *job shops*, o mais apropriado seria a adoção do BM para o acompanhamento das ordens com base no consumo do pulmão. No CONWIP, como padrão inicial do sistema, as ordens seguem o FCFS após entrarem no sistema. No entanto, pode-se aplicar outras regras de prioridade de acordo com as particularidades do sistema de produção.

Quanto ao dimensionamento do *buffer*, no DBR adota-se, normalmente, regras empíricas existentes na literatura, ao passo que, no CONWIP o mais comum é a análise do desempenho em relação a produtividade ou ao tempo de fluxo com base na Lei de Little. A liberação dos materiais para o sistema, no DBR, segue a programação da liberação, que é decorrente da programação do gargalo. No caso do CONWIP, a liberação ocorre respeitando o limite de WIP na linha (maxWIP).

Na literatura do DBR, o *Buffer Management* que serve como mecanismo de controle de produção e auxilia no ajuste do dimensionamento do *time buffer*. Na literatura do CONWIP, em contrapartida, há abordagens de *card controlling*, para ajustes na quantidade de cartões na linha de acordo com o desempenho do sistema.

Por fim, em relação ao ambiente, o DBR foi desenvolvido em uma perspectiva mais flexível, considerando ambientes com montagem e diversos roteiros. O CONWIP, originalmente, foi concebido para linhas simples como o *flow shop*, mas conta com variações propostas em estudos posteriores. O Quadro 5 resume as características de cada modelo.

Quadro 5 - Comparativo dos modelos DBR e CONWIP

	DBR	**CONWIP**
Foco do controle	Gargalo	Linha/segmento
Programação das ordens (*backlog list*)	Programação do gargalo com base no horizonte de programação, *time buffer* e regras de prioridade	Sequenciamento do *backlog list* com regras de prioridade
Sequenciamento nas máquinas (após a liberação)	*Status* do pulmão	FCFS
Dimensionamento do *buffer*	Predominantemente empírico (*rule of thumbs*)	Predominantemente focado em avaliações do desempenho com simulações (Lei de Little)
Liberação dos materiais	Programação da liberação dos materiais, de acordo com a programação do gargalo, que inclui o *time buffer*	De acordo com o limite de carga na linha (maxWIP)
Monitoramento do fluxo de produção	*Buffer Management*	*Card controlling*
Ambiente	Flexível: desenvolvido para ambientes com montagem e roteiros diversos	Restrito: desenvolvido, inicialmente, para *flow shops*

Fonte: Elaboração própria.

Para auxiliar o delineamento do presente estudo, os trabalhos comparativos do CONWIP e DBR apresentados no Quadro 2 (Capítulo 1) foram analisados quanto as características dos experimentos no Quadro 6. Observa-se que a grande maioria dos trabalhos constroem um ambiente produtivo para análise, apenas Gilland (2002) cita um caso real.

Entre os diversos cenários criados para avaliação dos modelos, verifica-se que os estudos consideraram tempos de processamento estocásticos, por meio de diversas distribuições de probabilidade. Em menor quantidade, há estudos que assumiram *setups* nas máquinas, possibilidade de falhas e reparos e produção de itens defeituosos. Na maior parte das simulações, as ordens chegam de forma dinâmica, conforme alguma distribuição de probabilidade para o intervalo entre chegadas.

Quanto ao controle da liberação dos *jobs* para a linha ao simular o DBR, praticamente todos os estudos controlaram por meio do WIP, ou seja, após o processamento de um *job* na máquina gargalo, pode haver a entrada de outra ordem no sistema. Isso torna as modelagens do CONWIP e DBR mais parecidas, diferindo apenas o ponto de controle na linha e omitindo a lógica de programação proposta inicialmente para o DBR.

Quadro 6 - Características dos estudos comparativos

Estudo	Características do experimento	Indicadores
Gilland (2002)	• *Flow shop;* • Baseado em um caso real; • Cenários com uma e duas restrições; • Variação dos tempos de processamento (exponencial); • Comparação de controles de liberação; • Controle da liberação do DBR pelo WIP.	• Utilização do gargalo; • Nível de WIP.
Kim; Davis e Cox III (2003)	• *Flow shop;* • Diferentes posições do gargalo; • Quebras nas máquinas em diferentes posições da linha; • Variação dos tempos de processamento (uniforme e exponencial); • Controle da liberação do DBR pelo WIP.	• *Output* médio; • WIP médio; • Atraso médio; • Quantidade média de pedidos atrasados.
Koh e Bulfin (2004)	• *Flow shop;* • Variação dos tempos de processamento (exponencial negativa); • Controle da liberação do DBR pelo WIP.	• *Output*; • Despesas operacionais; • Lucro líquido; • WIP; • Utilização;
Jodlbauer e Huber (2008)	• *Flow shop;* • Variação da demanda; • Variação dos tempos de processamento e *setups* (exponencial negativa); • Quebra de máquinas; • Taxa de produtos com defeito; • Não evidencia se usou controle de WIP ou *time buffer*.	• Nível de serviço; • WIP.
Fredendall; Ojha e Patterson (2010)	• *Job shop;* • Severidade do gargalo; • Diferentes níveis de capacidade protetiva dos recursos; • Taxa de chegada dos *jobs* (erlang e uniforme); • Variação dos tempos de processamento (gamma); • Diferentes níveis do WIP e formas de controlá-lo; • Controle da liberação do DBR pelo WIP.	• Percentual de atraso; • Tempo de fluxo; • *Lead time*; • *Jobs* no sistema; • Frequência de flutuação do gargalo.
Ng; Bernedixen e Syberfeldt (2012)	• *Flow shop;* • Diferentes posições do gargalo; • Variação dos tempos de processamento (log-normal); • Determinação da quantidade de *buffers* e a localização desses com otimização; • Controle da liberação do DBR pelo WIP.	• Vazão (*throughput*); • Tempo de fluxo.
Lee e Seo (2016)	• *Flow shop;* • Diferentes posições do gargalo; • Taxa de chegada dos *jobs* (Poisson); • Tempos de processamento constantes.	• WIP; • Nível de serviço.
Thürer et al. (2017b)	• *Flow shop;* • Diferentes posições do gargalo; • Severidade do gargalo; • Taxa de chegada dos *jobs* (exponencial); • Variação dos tempos de processamento (erlang truncada); • Diferentes regras de despacho; • Diferentes níveis do WIP; • Controle da liberação do DBR pelo WIP.	• Tempo de fluxo; • *Lead time*; • Percentual de atraso; • Atraso médio

Fonte: Elaboração própria.

Apesar de Framinan; González e Ruiz-Usano (2003) afirmarem que a maior parte dos estudos sobre CONWIP aborda o dimensionamento dos cartões, não se observou essa representatividade nos estudos comparativos. Somente Thürer et al. (2017b) testaram a variação dos níveis de WIP no sistema.

Dessa forma, continua pertinente a observação de Framinan; González e Ruiz-Usano (2003) quanto a importância de analisar o impacto das decisões, dimensionamento do WIP e sequenciamento da lista de liberação, conjuntamente no desempenho do CONWIP. Em todos os artigos, o sequenciamento dos *jobs* para liberação na linha seguiu regras de prioridade como urgência das ordens, data prometida, etc.

Em relação a diferenças entre a capacidade do gargalo e das demais máquinas, análise sugerida por Framinan; González e Ruiz-Usano (2003), apenas Fredendall; Ojha e Patterson (2010) e Thürer et al. (2017b), avaliaram diferentes níveis de severidades do gargalo. No entanto, nesse último estudo não houve representatividade desse fator nos resultados do experimento.

Alguns experimentos avaliaram diferentes posições do gargalo na linha. Kim; Davis e Cox III (2003) observaram que o CONWIP alcançou melhor resultado quando a restrição estava localizada no final da linha e o DBR teve bons resultados quando o gargalo estava no meio ou no final da linha.

Ng; Bernedixen e Syberfeldt (2012) observaram que independentemente da localização do gargalo na linha (início, meio ou final), o DBR superou o CONWIP e demais mecanismos puxados analisados. Já Thürer et al. (2017b) afirmaram que quando o gargalo era na primeira estação do *flow shop*, o DBR teve desempenho superior ao CONWIP, mas na última posição da linha os modelos foram semelhantes. Observa-se, assim, diferentes conclusões quanto ao impacto da posição, de acordo com o delineamento dos experimentos.

CAPÍTULO 5 – MÉTODO DE PESQUISA

Este capítulo apresenta o método de pesquisa. Inicialmente, é caracterizada a tipologia da pesquisa desenvolvida e, posteriormente, são descritas as etapas e técnicas aplicadas para a condução da mesma. Por fim, detalha-se o ambiente de produção em estudo.

5.1 Caracterização da pesquisa

A presente tese compara o desempenho dos mecanismos CONWIP e DBR em um *assembly flow shop* com produção MTO. Diante desse objetivo, esta pesquisa possui abordagem quantitativa, uma vez que considera um conjunto de variáveis com o intuito de identificar relações entre essas (BERTRAND; FRANSOO, 2002). Dentro da classe quantitativa, esta pode ser classificada como axiomática descritiva, pois analisa um problema idealizado com o objetivo de compreender e explicar as características do fenômeno modelado (MORABITO NETO; PUREZA, 2012).

Quanto aos procedimentos de investigação, a pesquisa classifica-se como experimental. A pesquisa experimental é caracterizada pelo desenvolvimento de experimentos em um sistema sob condições controladas, que permitem ao pesquisador a manipulação e o controle de variáveis a fim de observar as variações decorrentes dessas ações no fenômeno em estudo (MIGUEL, 2007; PRODANOV; FREITAS, 2013).

Para a condução dos experimentos, utilizou-se a Simulação por Eventos Discretos. A simulação é definida por Stewart (2004, p.04) como a “experimentação com uma imitação simplificada (em computador) de sistemas de operações à medida que esses evoluem no tempo com o objetivo de melhor compreendê-los e/ou melhorá-los”.

Para a validade da pesquisa axiomática com simulação, Bertrand e Fransoo (2002) destacam a importância dos seguintes elementos: (a) justificativa da escolha desse método e das hipóteses a serem testadas; (b) delineamento do experimento; (c) análise estatística dos resultados; (d) interpretação dos resultados.

Stewart (2004) justificam a utilização da simulação em virtude da natureza dos sistemas de operações, que envolvem variabilidade, componentes/operações interconectadas e complexidade combinatória (grande conjunto de possibilidades) ou dinâmica (diversas interações no sistema ao longo do tempo). A escolha pela simulação, então, está pautada na

complexidade do problema e no interesse em analisar a variabilidade e os efeitos a que estão sujeitos o ambiente de estudo (produção contra pedido em um sistema com operação de montagem e controle da liberação de ordens). Além disso, a simulação é o principal método aplicado em trabalhos que abordam essa problemática, como sinalizado por Jaegler et al. (2017).

As premissas assumidas no modelo, assim como, o delineamento dos experimentos foram norteados pela literatura pertinente ao tema, como sugerem Bertrand e Fransoo (2002) e detalhados nos tópicos seguintes. Os resultados do estudo foram analisados por meio da aplicação de ferramentas estatísticas (técnica analítica), a fim de investigar as relações de dependência entre as variáveis e a significância estatística dos resultados. Por fim, foram discutidos e interpretados os resultados, confrontando-os com estudos anteriores.

5.2 Etapas da pesquisa

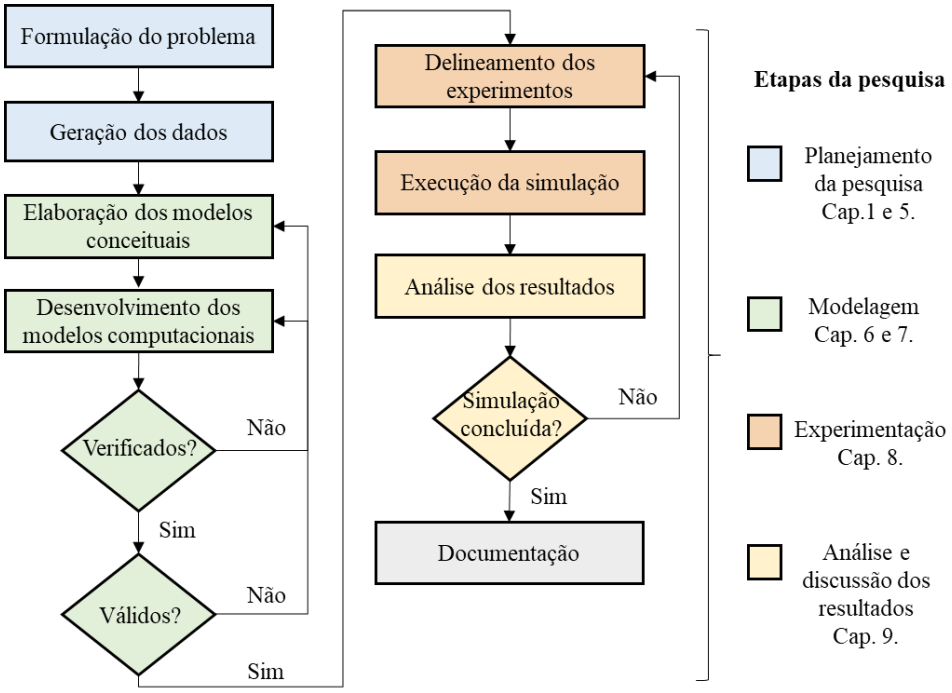
Para condução deste trabalho, seguiu-se quatro grandes etapas: (a) planejamento da pesquisa; (b) modelagem; (c) experimentação; (d) análise e discussão dos resultados. Essas etapas abrangem as fases de um projeto de simulação, que segundo Winston (2004) envolve:

a) Formulação e estruturação do problema: compreende a delimitação do problema a ser estudado, definição do objetivo e perguntas que nortearão o projeto. Na presente pesquisa, o detalhamento dessa primeira fase foi realizado no Capítulo 1;
b) Coleta de dados: como esta pesquisa é de natureza teórica, não houve coleta de dados em campo. Variáveis aleatórias foram geradas para caracterizar o ambiente estudado, como exposto na seção seguinte (5.3);
c) Desenvolvimento do modelo conceitual: a modelagem conceitual visa representar de forma simplificada os elementos essenciais do sistema em estudo. Como esta pesquisa objetiva simulação de dois sistemas, foram elaborados dois modelos. O modelo conceitual do CONWIP é descrito na seção 6.1 e o do DBR é detalhado no tópico 7.1;
d) Desenvolvimento do modelo computacional: refere-se à codificação do modelo conceitual, que neste estudo foi realizada no *software AnyLogic* versão *8.3.3 Professional*. As seções 6.2 e 7.2 detalham a implementação dos modelos CONWIP e DBR, respectivamente. O Apêndice A descreve os elementos do *software* empregados na construção dos modelos;

e) Verificação: análise do modelo de simulação a fim de identificar se o mesmo está funcionado corretamente, ou seja, não possui erros;
f) Validação: objetiva avaliar se o modelo representa de forma realista o sistema, atende ao propósito do estudo e se os resultados são confiáveis. As estratégias adotadas para verificação e validação dos modelos CONWIP e DBR são descritas nas seções 6.3 e 7.3, respectivamente;
g) Delineamento dos experimentos: abrange o planejamento dos experimentos, a identificação dos fatores, os níveis dos fatores e as variáveis de resposta que serão observadas. Como também compreende a determinação de parâmetros como duração da simulação, tempo de aquecimento e quantidade de replicações (Capítulo 8);
h) Execução do modelo computacional: consiste em "rodar" os diferentes experimentos no *software* de simulação;
i) Análise dos *outputs* do modelo: interpretação dos resultados dos modelos de simulação. Nesta pesquisa, analisou-se o desempenho dos sistemas CONWIP e DBR com suporte dos *softwares R* e *AnyLogic* em duas fases (estatística descritiva e Análise de Variância) apresentadas no Capítulo 9;
j) Documentação e implantação dos resultados: compreende a construção dos relatórios do projeto e, quando se refere a um problema aplicado, realiza-se a implantação das proposições. Nesta pesquisa, entende-se que a documentação é o detalhamento feito ao longo dos Capítulos e Apêndices do presente documento.

A Figura 13 apresenta o encadeamento das fases do projeto de simulação, as seções e capítulos que detalham cada uma das quatro etapas desta pesquisa. A seção seguinte deste Capítulo finaliza a primeira etapa (planejamento da pesquisa).

Figura 13 - Etapas da pesquisa



Fonte: Adaptado de Winston (2004).

5.3 Caracterização do ambiente de produção

O ambiente de produção em estudo é um sistema de montagem (*assembly flow shop*), em resposta as oportunidades de pesquisa evidenciadas nos Quadros 1 e 2 do Capítulo 1. Nesse ambiente, todos os *jobs* seguem o mesmo roteiro, de forma que as linhas de fabricação e finalização são análogas a um *flow shop* puro (PINEDO, 2012).

Como esta pesquisa é um dos primeiros trabalhos comparativos sobre os sistemas CONWIP e DBR em ambientes com fluxos convergentes, optou-se por uma configuração simples: apenas uma operação de montagem, abastecida por duas linhas de fabricação com três máquinas e seguida por uma linha de finalização com duas operações. Cada operação possui apenas uma máquina (capacidade = 1 *job*), sendo o ambiente composto por nove máquinas. Essa configuração é semelhante a um dos casos analisados por Khojasteh-Ghamari (2009) na comparação de sistemas de controle da produção.

A produção nesse ambiente é *make-to-order*, não há estoques de produtos acabados e o processo de fabricação inicia-se somente após o recebimento de um pedido. A produção MTO foi caracterizada pela variabilidade dos tempos de processamento e da chegada dos *jobs*, assim como representado por Thürer e Stevenson (2018).

A chegada dos pedidos no sistema ocorre de forma dinâmica, caracterizada pela taxa de chegada segundo uma distribuição de probabilidade exponencial, como adotado por Rocha et al. (2015) e Thürer e Stevenson (2018). Na simulação, considerou-se a taxa de 9 *jobs*/unidade de tempo. Cada pedido recebido é composto por duas ordens de fabricação, que após a montagem são unificadas. Os pedidos possuem uma data prometida para entrega definida na chegada, essa data é um atributo comum as ordens de fabricação.

A data prometida gerada para determinada ordem i (DD_i) é dada em função da soma do momento de chegada do *job* no sistema (a_i) e um fator de tolerância dado por uma distribuição uniforme (24,0; 30,0). Esse fator de tolerância foi determinado de forma experimental, de acordo com tempos médios de processamento e tempos com filas, de modo semelhante a Thürer e Stevenson (2018), e um percentual médio de ordens atrasadas inferior a 10%. Após a calibração, o mesmo valor do fator foi adotado para a simulação dos modelos CONWIP e DBR a fim de garantir condições semelhantes para comparação.

Os tempos de processamento das ordens nas máquinas seguem uma distribuição de probabilidade Erlang, visto que possui ampla aplicabilidade na geração de variáveis aleatórias contínuas e positivas e é especialmente usada na predição de tempos de espera em sistemas de filas. A distribuição Erlang também já foi utilizada para representar tempos de processamento em estudos de simulação sobre a produção MTO (BEEMSTERBOER et al., 2017; THÜRER et al., 2017b).

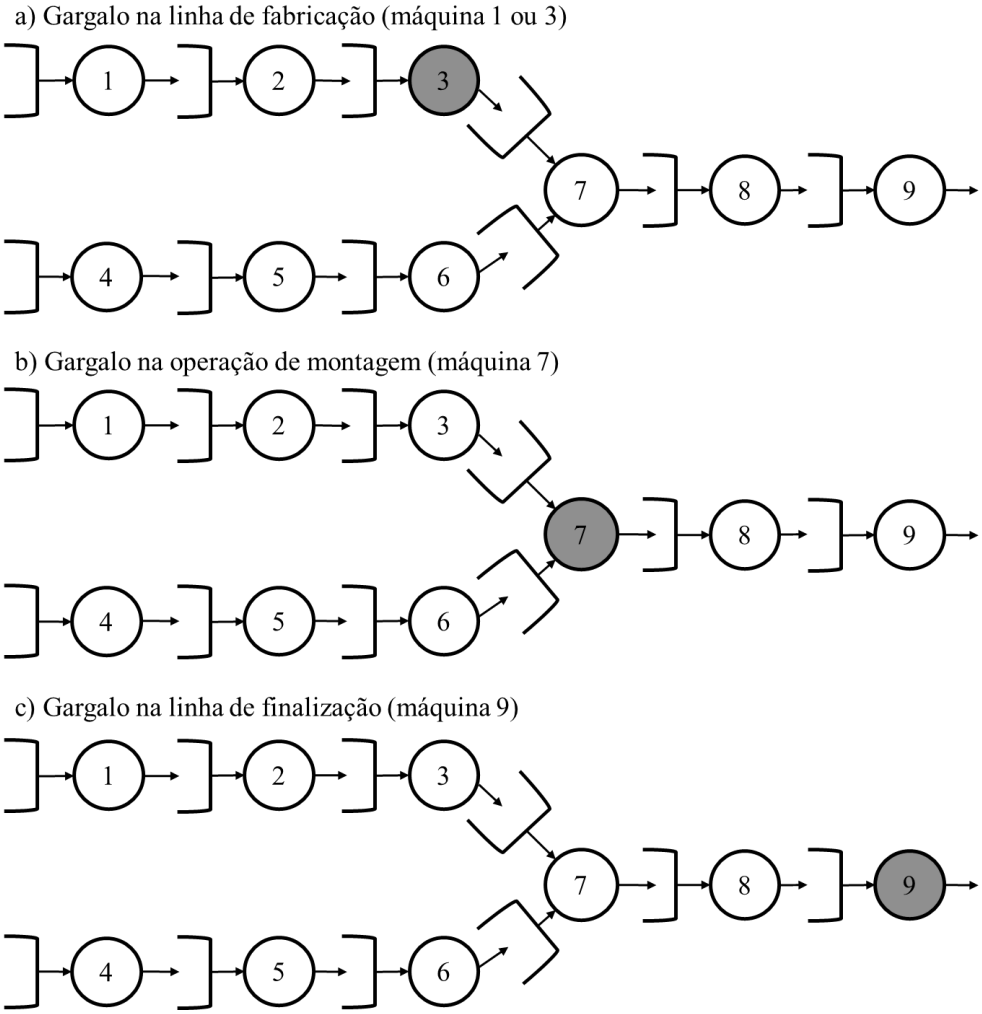
Adotou-se os seguintes parâmetros da distribuição Erlang: k = 4 (quantidade de exponenciais somadas), p = uniforme (0,7; 1,3) (média dos tempos de processamento) e a média das variáveis exponenciais igual a p/k. Considerou-se que a média dos tempos nas operações (p) variou conforme uma distribuição de probabilidade uniforme, a fim de caracterizar diferentes tamanhos das ordens de produção.

De modo semelhante a outros trabalhos teóricos sobre o DBR (THÜRER; STEVENSON, 2017, 2018), foram considerados níveis de severidade do gargalo. Para caracterizar o desbalanceamento, o tempo médio de processamento dos recursos não críticos foi reduzido em relação ao gargalo, de acordo com três níveis de severidade: (a) baixa (operação gargalo 10% mais lenta); (b) média (gargalo 30% mais lento) e, (c) alta (gargalo 50% mais lento).

Diferentes posições do recurso crítico foram avaliadas no estudo, como mostra a Figura 14. Considerou-se duas posições na linha de fabricação, uma no início e outra no final da linha (recursos 1 ou 3, Figura 14a), o gargalo na operação de montagem (máquina 7, Figura 14b) e uma posição na linha de finalização (máquina 9, Figura 14c).

Dessa forma, admitiu-se que o impacto do gargalo nas estações 4 ou 6 no desempenho do sistema é semelhante ao das posições 1 ou 3, respectivamente. Para reduzir a complexidade computacional do problema, optou-se por quatro posições que representasse cenários distintos (início, meio e final do processo de produtivo) e compreendessem as três áreas do sistema de montagem (fabricação, montagem e finalização) na simulação. Em todos os casos, o caminho crítico, sequência de operações com maior tempo médio total, compreende as estações 1-2-3-7-8-9.

Figura 14 - *Assembly flow shop* em estudo com diferentes posições da estação gargalo



Fonte: Elaboração própria.

A Tabela 1 apresenta a síntese das características do ambiente simulado. O desempenho desse ambiente foi avaliado com base em indicadores de desempenho que remetem ao cumprimento do prazo, velocidade de fluxo e utilização da capacidade.

Tabela 1 - Síntese das características do ambiente simulado

Parâmetros	**Dados**
Ambiente de produção	*Assembly flow shop*
Quantidade de operações	9
Taxa de chegada	0,9 *jobs*/unidade de tempo (Exponencial)
Média dos tempos de processamento (variação do tamanho das ordens)	*p* = Uniforme (0,7; 1,3)
Tempos de processamento	4-Erlang (*p*)
Severidade do gargalo	10%, 30% ou 50%
Posição do gargalo	Operação 1, 3, 7 ou 9
Caminho crítico	Operações 1, 2, 3, 7, 8 e 9
Fator de tolerância para definição da data prometida	Uniforme (24,0; 30,0)

Fonte: Elaboração própria.

Na simulação dos sistemas, assumiu-se que todos os *jobs* são aceitos para o processamento e aguardam a liberação no *pre-shop pool* para controle do estoque em processo, conforme o modelo (CONWIP ou DBR). Além disso, admitiu-se que todos os insumos para a fabricação estavam disponíveis e não foram consideradas outras variabilidades no processo como quebras e paradas das máquinas.

CAPÍTULO 6 – MODELAGEM CONWIP

Este capítulo detalha a modelagem do sistema CONWIP. Inicialmente, caracteriza-se o modelo conceitual, a configuração do controle da liberação e o sequenciamento dos *jobs*. Em seguida, descreve-se a implementação computacional e as estratégias adotadas para verificação e validação do modelo CONWIP.

6.1 Modelo conceitual CONWIP

Os *inputs* do modelo CONWIP são: (a) taxa de chegada dos *jobs*; (b) parâmetros para definição dos tempos de operação; (c) fator de tolerância para definição do prazo de entrega; (d) regra de sequenciamento do *pre-shop pool*; (e) *loops* e quantidade de cartões por segmento.

O modelo compreende cinco processos principais: chegada, sequenciamento dos *jobs*, controle da liberação, processamento, saída dos *jobs* e cálculo dos indicadores. O processo de chegada refere-se à geração dos *jobs*, sorteio da média dos tempos de processamento e determinação da data prometida para entrega, conforme parâmetros da Tabela 1.

Em seguida, os pedidos entram na fila do *pre-shop pool*, onde são sequenciados. Observa-se que na literatura sobre produção MTO, a maioria dos estudos consideram regras simples para o sequenciamento da fila de entrada e das filas das máquinas, em virtude da dinâmica do ambiente de produção contra pedido. Diante disso, neste trabalho aplica-se a regra *Earliest Due Date* (EDD) para o sequenciamento do *backlog list* (LEU, 2000; THÜRER et al., 2017a).

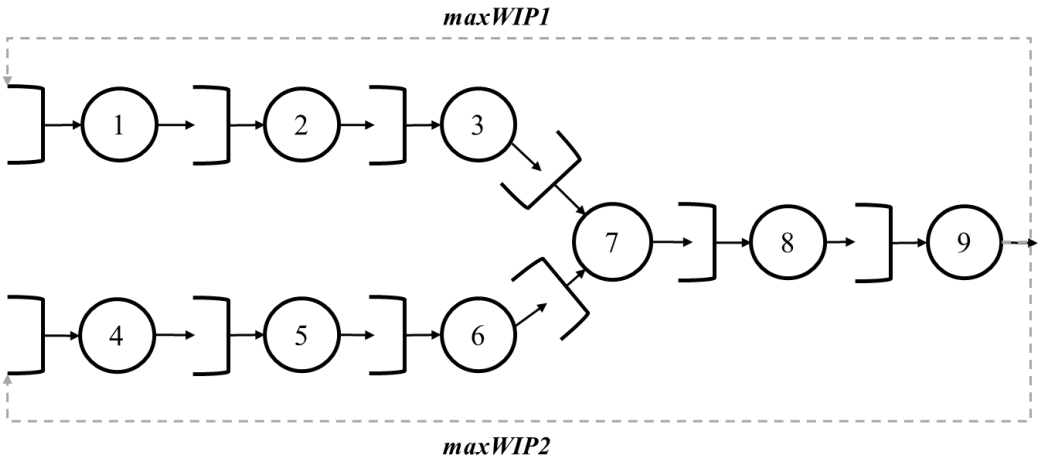
Na regra EDD, os *jobs* com menor data prometida, ou seja, menor folga para a entrega, são priorizados na liberação. Além de ser uma regra comumente utilizada na literatura sobre o CONWIP, apresentou bom desempenho em ambientes com montagem (FRY et al., 1989; LU; HUANG; YANG, 2011) e prioriza o atendimento no prazo, critério competitivo relevante para a estratégia MTO.

Como mencionado na seção 5.3, cada pedido é composto por duas ordens, que seguem em paralelo nas linhas de fabricação. O momento para liberação das ordens de fabricação é definido pela disponibilidade de cartões no *loop*. Na literatura sobre o CONWIP, observam-se diversas configurações dos *loops* para controle da liberação. Neste estudo, buscou-se configurar o CONWIP de forma mais parecida com a proposta inicial do modelo, de modo que o controle de WIP não dependesse da posição do gargalo.

Logo, o CONWIP foi configurado com dois *loops* (segmento 1 – máquinas 1, 2, 3, 7, 8, e 9; segmento 2 – máquinas 4, 5, 6, 7, 8 e 9) para controle de estoque em processo (maxWIP1 e maxWIP2), como mostra a Figura 15. Essa configuração também foi analisada por Huang et al. (2015) em ambiente com montagem e produção de alta variedade e pequenos lotes, com bons resultados em termos de tempo de fluxo.

Os limites de WIP em cada segmento (maxWIP1 e maxWIP2) podem ser iguais ou diferentes. Inicialmente, considerou-se valores diferentes, uma vez que Hopp e Spearman (2013) sugerem que para o segmento com maior tempo total (a linha crítica) haveria um maxWIP maior.

Figura 15 - Configuração dos *loops* e limites de WIP no modelo CONWIP



Fonte: Elaboração própria.

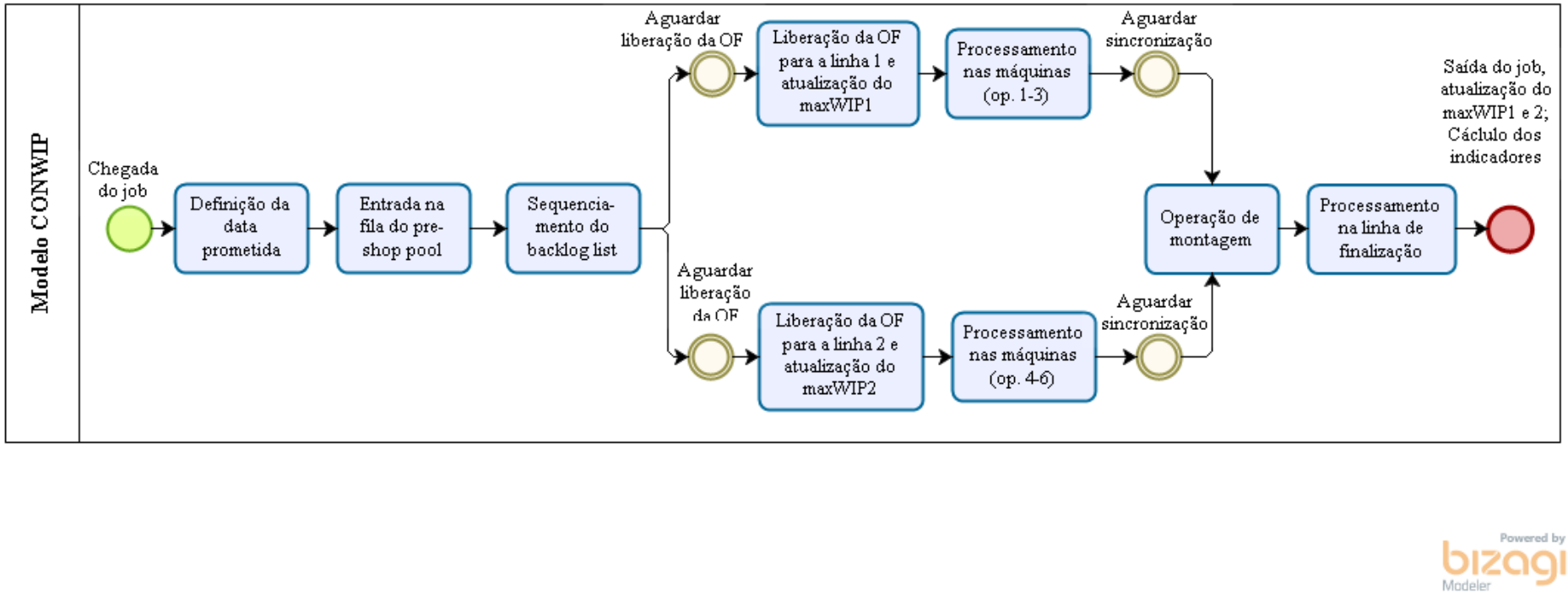
Os cartões foram segmentados por *loop*, como sugerido por Khojasteh-Ghamari (2009). Ao liberar um pedido que aguardava na fila de entrada, há o incremento dos níveis de WIP em uma unidade para cada segmento e são utilizados dois cartões (um para cada componente a ser processado nas linhas de fabricação).

Na operação de montagem (máquina 7), as ordens de fabricação dos componentes são unidas e prossegue apenas uma ordem de produção unificada na linha de finalização. Nesse momento, os dois cartões, que acompanhavam os componentes, seguem com o produto em processamento. Ao concluir a produção do pedido na operação 9, os dois cartões são liberados para as respectivas linhas de fabricação, permitindo a entrada de novas ordens de fabricação. Como simplificação, nesta pesquisa não foi considerado o controle dinâmico da quantidade de cartões (*card controlling*).

Após liberadas no *pre-shop pool*, as ordens seguem a sequência definida na entrada, ou seja, a ordenação das filas das máquinas é FCFS. Essa regra foi adotada uma vez que todos *jobs* seguem o mesmo roteiro no sistema de produção (*assembly flow shop*).

Com a saída de um *job*, são atualizados os níveis de WIP de cada *loop* e calculados os indicadores do pedido (tempo de fluxo, *lead time* e atraso). Ao final de cada simulação, são calculados os *outputs* do modelo: tempo médio de fluxo das ordens, *lead time* médio, utilização do gargalo, percentual de ordens atrasadas e atraso médio para avaliação do sistema. A Figura 16 apresenta o modelo conceitual do CONWIP, o termo “OF” refere-se a ordem de fabricação.

Figura 16 - Modelo conceitual da simulação do sistema CONWIP



Fonte: Elaboração própria.

6.2 Modelo computacional CONWIP

O modelo conceitual do CONWIP foi implementado no *AnyLogic* versão *Professional 8.3.3*, com emprego dos blocos da aba de Modelagem de Processos (*Process Modeling Library*) do *software* descritos no Apêndice A e a entrada dos dados do ambiente em estudo (seção 5.3). A Figura 17 apresenta a captura da tela principal do modelo de simulação do CONWIP no *software*.

O primeiro elemento do modelo é a "chegada", onde são criados os *jobs* (agentes do modelo). No momento da chegada, os pedidos recebem como atributos: (a) identificação/número do pedido; (b) data prometida de entrega; (c) tempo médio de processamento, que indica o tamanho do *job*.

Após a entrada dos *jobs* no sistema, foi introduzido o elemento "*copia*", que cria as ordens de fabricação (duplica as entidades). Esse artifício não gasta tempo no modelo e reproduz todos os atributos definidos anteriormente para a entidade cópia. As ordens aguardam a liberação no *pre-shop pool* da respectiva linha de fabricação (*preshop1* e *preshop2*), onde são ordenadas de acordo com a regra de sequenciamento EDD.

Os elementos "*control1*" e "*control2*" executam o controle da liberação com base nos parâmetros maxWIP1 e maxWIP2 e nas variáveis WIP1 e WIP2, respectivamente. O modelo monitora a quantidade de ordens que estão em processamento, entre a "*fila1*" e a "*saida*" (WIP1) e entre a "*fila4*" e a "*saida*" (WIP2). Se o maxWIP1 for igual ao maxWIP2, então as duas ordens de fabricação são liberadas simultaneamente.

Depois de liberadas, as ordens de fabricação seguem para o processamento nas máquinas nas respectivas linhas. Cada máquina é antecedida por uma fila que não possui restrição de capacidade. Concluída a fabricação do componente, esses aguardam nas filas (*fila7f1* e *fila7f2*).

Para garantir que a operação de montagem faça a união dos componentes de um mesmo pedido, ou seja, as ordens de fabricação tenham a mesma numeração, utilizou-se os elementos "*comp1*" e "*comp2*". Esses blocos, juntamente com comandos lógicos inseridos na "*fila7f1*" e "*fila7f2*", verificam a numeração das ordens que chegam nas filas, ou seja, fazem a sincronização da montagem. Caso identifique os dois componentes de uma ordem, esse são enviados para a "*fila7comp1*" e "*fila7comp2*".

A operação de montagem foi modelada com dois elementos: o "*comb*", para fazer a união das ordens, um artifício que também não gasta tempo da simulação; o "*m7*", que representa o

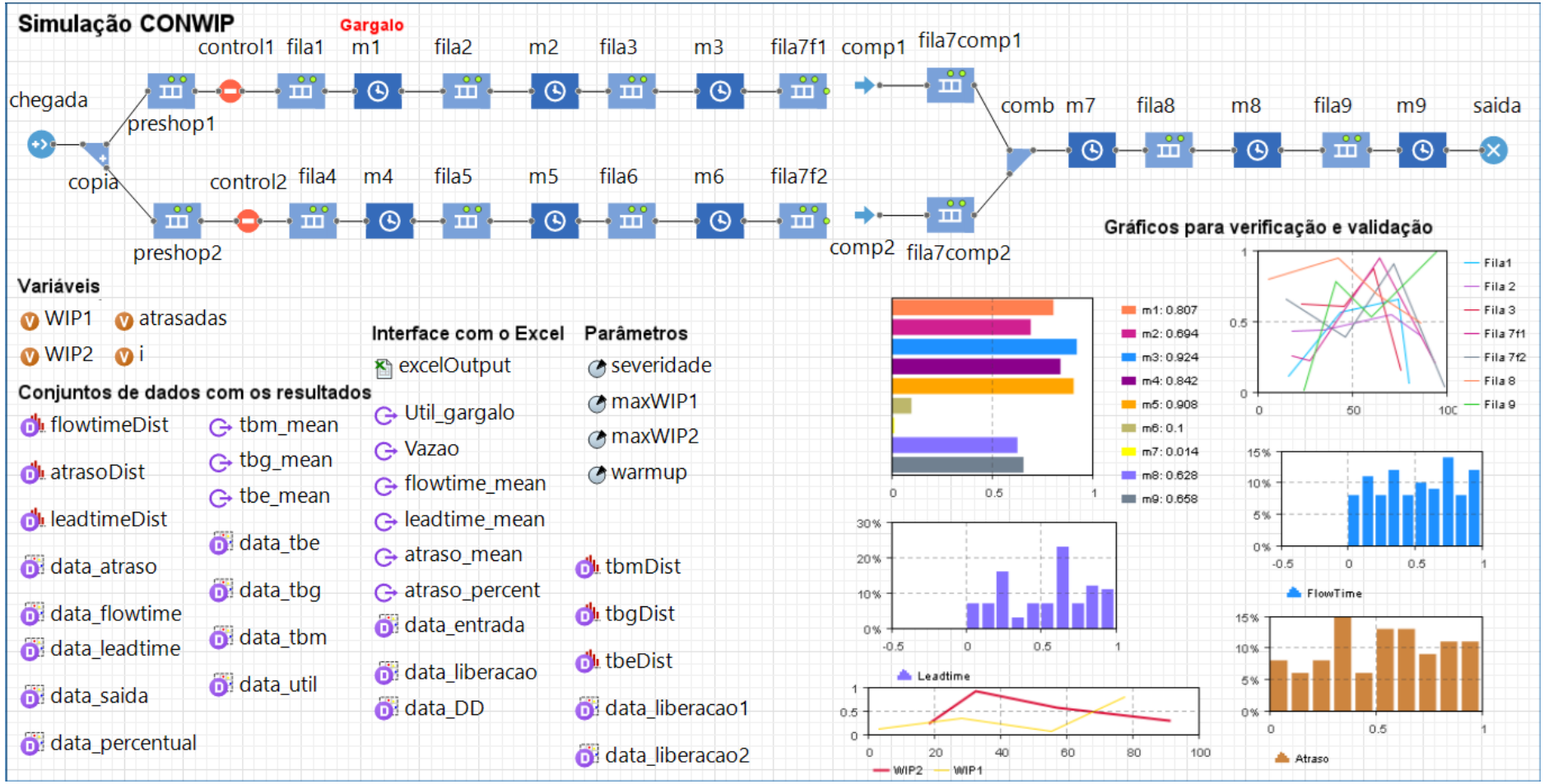
tempo de processamento da operação. Em seguida há a finalização do *job* nas máquinas *m8* e *m9*.

Com a conclusão de um *job* na última máquina (*m9*), os níveis de WIP (variáveis WIP1 e WIP2) são atualizados e permite-se a entrada de uma nova ordem de fabricação nas linhas de fabricação. No objeto "*saida*", a entidade é destruída e são armazenadas as informações quanto aos indicadores tempo de fluxo, *lead time*, atraso e percentual de ordens atrasadas.

Na Figura 17, abaixo do sistema de montagem (blocos lógicos) estão ícones referentes a variáveis como os níveis de WIP das linhas e "*i*", que auxilia na numeração das ordens. Há também parâmetros do modelo, relacionados ao limite de WIP (maxWIP1 e maxWIP2) e à severidade do gargalo, etc.

Diversas listas de dados foram utilizadas para armazenar resultados da simulação, como o atraso (*data_atraso*), *lead time* (*data_leadtime*) e o tempo de fluxo das ordens (*data_flowtime*), etc. Além disso, foram criados indicadores para guardar os resultados finais das variáveis de resposta e elaborados gráficos como histogramas e de barras com a utilização dos recursos. Esses elementos também auxiliaram o processo de verificação e validação do modelo. Toda a configuração dos blocos do modelo de simulação para representar a lógica do CONWIP no *AnyLogic* é feita com comandos em *Java* (linguagem de programação orientada a objetos).

Figura 17 - Captura da tela principal do modelo de simulação do CONWIP



Fonte: Elaboração própria.

6.3 Verificação e validação CONWIP

Para verificação e validação dos modelos computacionais, tanto do CONWIP como do DBR, foram aplicadas algumas técnicas sugeridas por Sargent (2013), tais como:

a) Animação: foi realizada a verificação visual das animações que representam o sistema de montagem, observação do fluxo dos materiais ao longo das linhas e do controle da liberação de novas ordens no sistema;
b) Gráficos de desempenho: foram criados gráficos e indicadores de desempenho para representar o comportamento do sistema ao longo da simulação, como o nível de WIP no sistema, tamanho das filas das máquinas, taxa de utilização dos recursos, histogramas dos atrasos, tempos de fluxo e *lead time* dos *jobs*;
c) Relacionamento entre os dados: foi avaliado se o modelo computacional representava corretamente a relação entre os dados. Dessa forma, analisou-se o impacto de variações do *buffer* (maxWIP) no tempo de fluxo das ordens, mudanças no fator de tolerância (*f*) da data prometida e o impacto no percentual de atrasos, etc.;
d) Testes degenerativos: observou-se o impacto do emprego de valores para *inputs* que comprometem o desempenho do sistema continuamente, como taxa de chegada dos *jobs* maior que a taxa de atendimento do sistema, em indicadores das filas e do WIP;
e) *Trace* (depurador): utilizou-se o comando "*traceln*" nos modelos de simulação para acompanhar todas as etapas de geração, definição de atributos (identificação da ordem, data prometida, tempo médio de processamento, etc.), processamento nas operações de fabricação, montagem, conclusão dos *jobs* e cálculo dos indicadores. Esse acompanhamento dos eventos com a impressão no *console* permitiu avaliar se a lógica do modelo de simulação estava correta e se a implementação foi adequada.

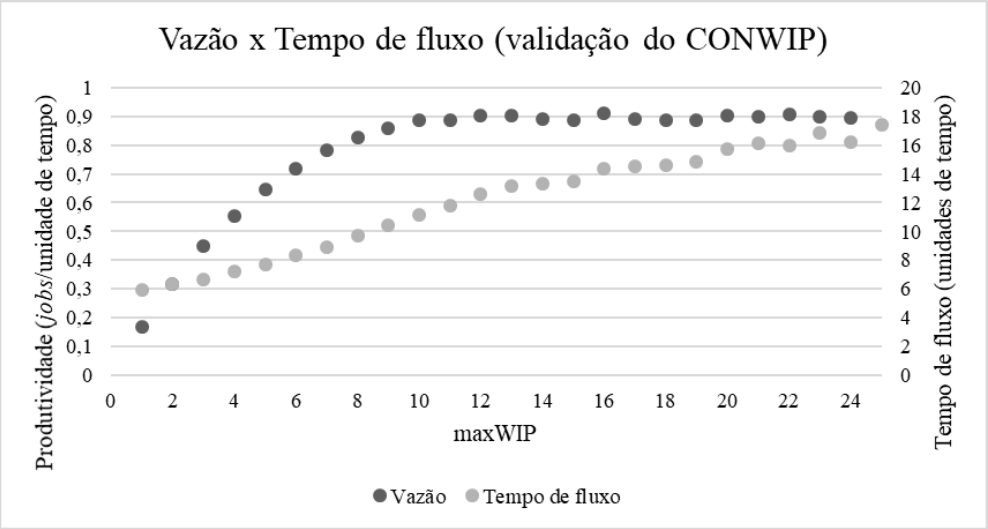
Além das inspeções visuais do modelo, Stewart (2004) e Law e Kelton (2000) também destacam a importância de analisar os resultados da simulação e compará-los com os resultados esperados. Nesse sentido, o modelo rodou com chegadas em intervalos regulares e tempos de processamento constantes. Os *outputs* do modelo foram impressos em planilhas do *Excel*. Esse registro permitiu averiguar os atributos das entidades, observar que a prioridade da regra de sequenciamento é seguida e os indicadores de desempenho são medidos de forma correta.

A Figura 18 apresenta os resultados de um dos testes realizados no processo de validação do modelo de simulação do CONWIP, quanto ao relacionamento dos dados (item c). Assumiu-se a taxa de chegada de 0,9 *jobs*/unidade de tempo, o tempo médio de processamento nas

máquinas dado por 4-Erlang(*p*), sendo *p*= uniforme (0,7; 1,3), com severidade baixa do gargalo (10%), como exposto na Tabela 1. Nesse teste, o limite de WIP foi considerado o mesmo para os dois segmentos, ou seja, maxWIP = maxWIP1 = maxWIP2. O limite variou de 1 a 25 *jobs* no sistema. Cada cenário foi simulado uma única vez, com duração de 10.000 unidades de tempo.

Como sugerido pela Lei de Little, a Figura 18 mostra o relacionamento entre as variáveis tempo médio de fluxo e vazão (produtividade) em relação as variações do parâmetro maxWIP. Inicialmente, um aumento no nível de WIP gerou um incremento da vazão do sistema e do tempo médio de fluxo das ordens. Posteriormente, a vazão começou a estabilizar em torno de 0,9 *jobs*/unidade de tempo (taxa de chegada do sistema), enquanto que o tempo de fluxo continuou aumentando. Esse resultado está coerente com o expresso na literatura do CONWIP (HOPP; SPEARMAN, 2013).

Figura 18 – Validação do CONWIP



Fonte: Elaboração própria.

CAPÍTULO 7 – MODELAGEM DBR

Este capítulo detalha a modelagem do sistema de DBR. Inicialmente, caracteriza-se o modelo conceitual, especialmente, a configuração do controle e programação dos *jobs*. Em seguida, descrevem-se a implementação do modelo computacional e as estratégias para verificação e validação.

7.1 Modelo conceitual DBR

De forma semelhante ao exposto no Capítulo 6 sobre o CONWIP, a simulação do DBR tem por objetivo representar a aplicação desse sistema de programação e controle da liberação de ordens no ambiente de montagem com produção contra pedido em estudo.

Os *inputs* do modelo DBR são: (a) taxa de chegada dos *jobs*; (b) parâmetros para os tempos de atendimento; (c) fator de tolerância para a data prometida; (d) lógica para programação do *backlog list*; (e) *time buffers*.

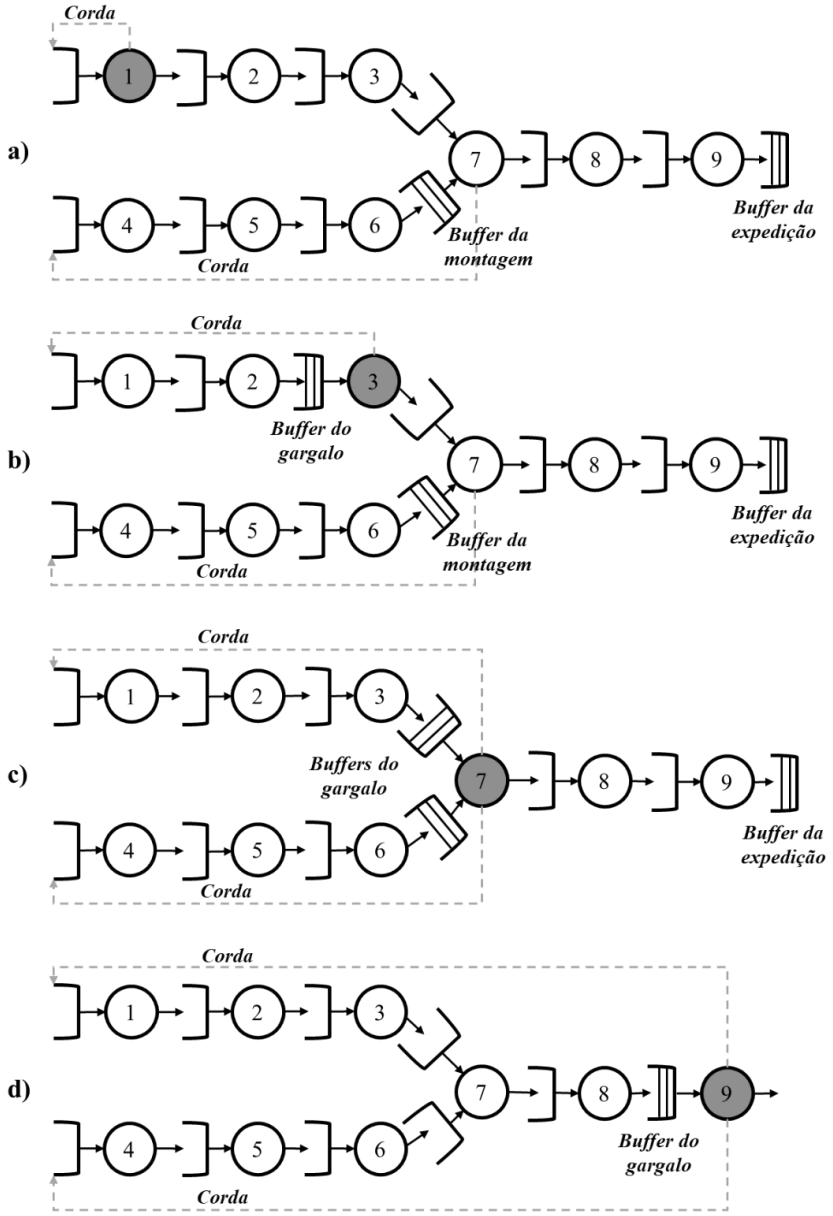
O modelo conceitual do DBR compreende as etapas de chegada, programação das ordens, controle da liberação, processamento e saída dos *jobs*/cálculo dos indicadores. O processo de chegada é idêntico ao do CONWIP, os *jobs* são gerados de acordo com a taxa de chegada, sorteia-se a média dos tempos de processamento e calcula-se o prazo para entrega (como descrito na Tabela1).

Depois disso, os *jobs* seguem para o *pre-shop pool* onde aguardam o momento para liberação. As principais diferenças entre as lógicas dos modelos estão na segunda e terceira etapa. Como mencionado no Capítulo 3, o DBR, originalmente, considera três tipos de *time buffer*: (a) do gargalo, que tem a função de proteger programação da restrição; (b) da montagem, empregado para evitar que as peças oriundas do gargalo aguardem por peças da linha não-crítica para sincronização na montagem; (c) de expedição, utilizado evitar que eventuais variabilidades nas máquinas após o gargalo comprometam o atendimento do prazo de entrega.

Como o estudo avalia diferentes posições do gargalo, em cada análise pode haver a consideração de diferentes *time buffers*, além disso, o *buffer* do gargalo pode assumir localizações distintas. Essas variações impactam na programação e controle da liberação das ordens no DBR. A Figura 19 apresenta as quatro possíveis configurações do DBR de acordo com as posições do gargalo analisadas nesta pesquisa:

a) Gargalo na primeira máquina do sistema: já que não há restrição de suprimento (seção 5.3), não há consideração do *time buffer* do gargalo, apenas os da montagem e expedição;
b) Gargalo na linha de fabricação (máquinas 3): nesse caso, são considerados três tipos de *time buffer* (gargalo, montagem e expedição);
c) Gargalo na montagem (máquina 7): nessa condição, aplica-se o *time buffer* do gargalo e o da expedição;
d) Gargalo na última máquina do sistema: como as ordens são imediatamente entregues após a conclusão da última operação (seção 5.3), não há necessidade do pulmão de expedição, supõe-se apenas o *time buffer* do gargalo.

Figura 19 - Configurações do DBR de acordo com a posição do gargalo no ambiente em estudo



Fonte: Elaboração própria.

Para a programação dos *jobs*, adotou-se a lógica apresentada na literatura clássica de DBR (COX III; SPENCER, 2002; CHAKRAVORTY; ATWATER, 2005). No entanto, diferentemente do exposto em Simons Jr. e Simpson III (1997), não foi estimado um horizonte para a programação. Diante da dinâmica da produção MTO, discutida na literatura que avalia a adoção do DBR nessa estratégia, optou-se por sequenciar os *jobs* à medida que esses são gerados, definindo o *Planned Release Date* (PRD). As ordens são priorizadas segundo o menor valor do momento planejado para a liberação (PRD).

De acordo com a posição do gargalo e os *time buffers*, o PRD é calculado de diferentes formas. Quando o gargalo está localizado na linha de fabricação (máquinas 1 ou 3), calcula-se o momento de liberação na linha crítica (PRD', estação 1) e na linha não crítica (PRD", máquina 4). Para a linha crítica, deduz-se da data prometida do *job* (DD_i), o *time buffer* da expedição (tbe), o tempo médio de processamento no gargalo (p_g) e, no caso da máquina 3, subtrai-se também o pulmão do gargalo (tbg).

Para cálculo do momento planejado para liberação na linha não crítica, adotou-se um *time buffer* da montagem expandido (tbm), que compreende todas as operações da linha de fabricação não crítica (máquinas 4-5-6), a montagem e as operações de finalização. Essa estratégia foi empregada a fim de simplificar a estimativa para a linha não-crítica, já que nos dois casos (gargalo na máquina 1 ou 3), o PRD" é calculado deduzindo da data de entrega apenas o tbm. Esse *time buffer* protege a sincronização da montagem e a entrega no prazo.

As Equações 7.1 e 7.2 apresentam as fórmulas para o caso do gargalo na primeira máquina, enquanto que as Equações 7.3 e 7.4 são as expressões para o cálculo da liberação no caso da terceira máquina.

Quando g = 1,

$$PRD' = DD_i - tbe - p_g \tag{7.1}$$

$$PRD'' = DD_i - tbm \tag{7.2}$$

Quando g = 3,

$$PRD' = DD_i - tbe - p_g - tbg \tag{7.3}$$

$$PRD'' = DD_i - tbm \tag{7.4}$$

Nos demais casos da posição do gargalo, operação de montagem (g = 7) ou última máquina do da finalização (g = 9), calcula-se apenas um momento planejado de liberação das ordens para ambas as linhas de fabricação (PRD'). Nesses cenários, todos os componentes oriundos das linhas de fabricação abastecem, ou seja, são processados no recurso gargalo.

Quando o gargalo estiver na máquina 7, subtrai-se da data de entrega do *job* (DD_i) o *time buffer* da expedição (tbe), o tempo médio de processamento do gargalo (p_g) e o *time buffer* do gargalo (tbg), como mostra a Equação 7.5. Quando a restrição for a última máquina do sistema, desconta-se da data prometida do *job* (DD_i) apenas o tempo de operação do gargalo (p_g) e o pulmão da restrição (tbg) como exposto na Equação 7.6.

Quando g = 7,

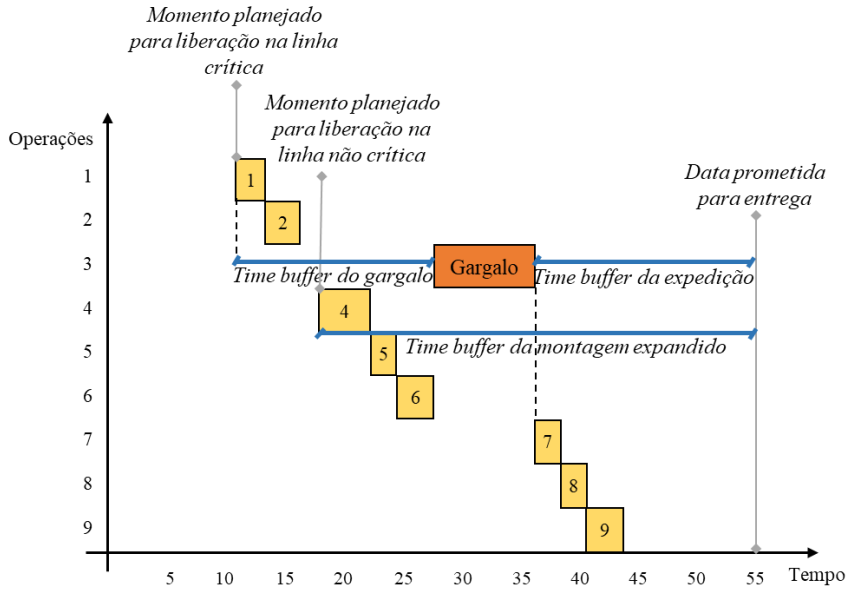
$$PRD' = DD_i - tbe - p_g - tbg \tag{7.5}$$

Quando g = 9,

$$PRD' = DD_i - p_g - tbg \tag{7.6}$$

A Figura 20 exemplifica a lógica da programação para trás realizada no DBR quando o gargalo está na máquina 3 (o exemplo é apenas ilustrativo, não houve estimação dos tempos). Observa-se que a consideração dos *time buffers* implica em realizar com certa antecedência as operações a fim de evitar que variabilidades comprometam a utilização do gargalo e a entrega no prazo.

Figura 20 – Lógica da programação e controle da liberação com o DBR com o gargalo em diferentes posições



Fonte: Elaboração própria.

Apenas a determinação dos PRDs, no entanto, não leva em consideração a carga do sistema para liberação dos *jobs*. Além disso, devido às variações dos tempos de processamento (processo estocástico), pode ocorrer dos *jobs* serem processados mais rápido do que o previsto

e o gargalo ficar disponível para processamento, enquanto há ordens aguardando a liberação no *pre-shop pool*. Ou ainda, os *jobs* podem demorar mais do que o planejado no processamento, de modo que, a liberação de um pedido no momento previsto faça com que a ordem aguarde por mais tempo no chão de fábrica, aumentando o WIP e o tempo de fluxo.

Diante disso, optou-se por considerar uma variável chamada *backlog* do gargalo (BLG), que indica a carga de trabalho (aproximada) que foi liberada para o recurso crítico, mas ainda não foi processada nesse recurso. Essa variável é calculada com base na soma dos tempos médios de processamento no gargalo das ordens que estão entre a fila da primeira máquina da linha de fabricação e a fila que antecede o recurso crítico.

Essa estratégia foi adotada com o objetivo de melhor representar o "tambor" na simulação do DBR, levando em consideração a carga imposta a restrição e emulando uma programação finita para frente (SIMONS JR.; SIMPSON III, 1997; SOUZA, 2005). Os valores de PRD indicam a prioridade das ordens o sequenciamento da fila de entrada e o controle da liberação dos *jobs* para o chão de fábrica é, então, condicionado a estimativa da carga de trabalho no gargalo.

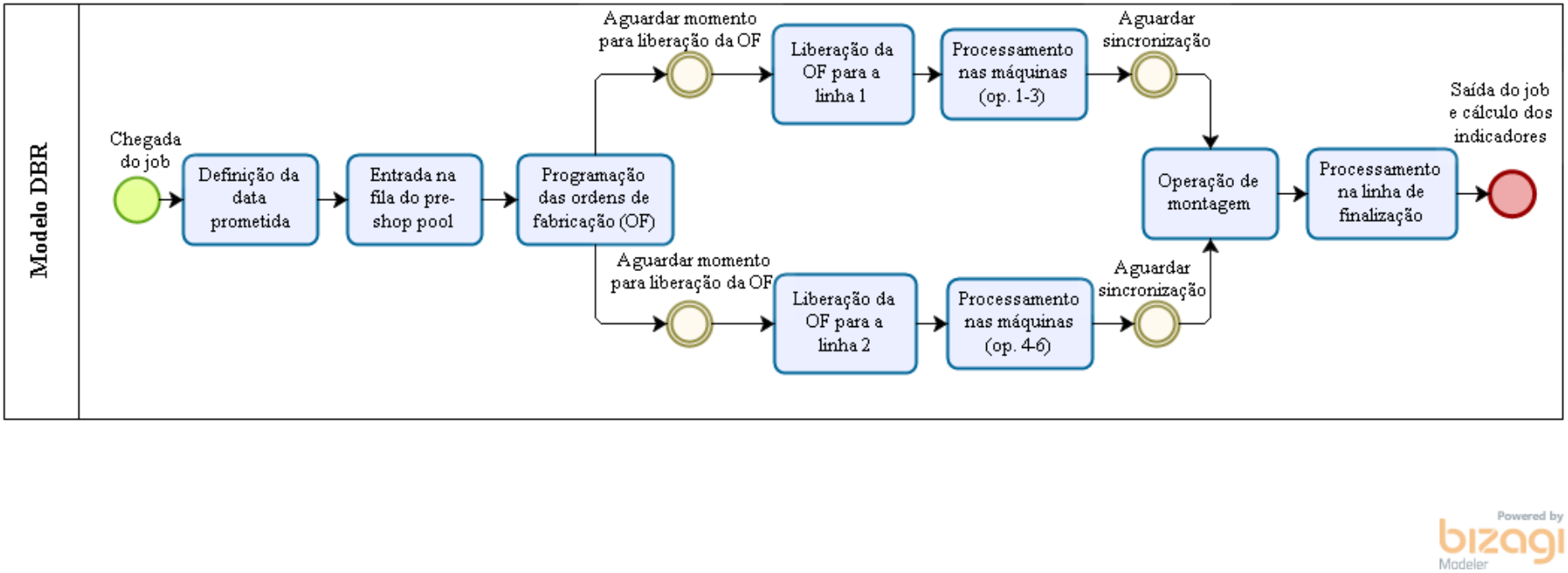
A liberação é dada em função da chegada do pedido (a_i), mais o BLG (estimativa do tempo que levará para a ordem ser atendida no gargalo), menos o *time buffer* do gargalo (estimativa do tempo que a ordem levará para processamento nas máquinas que antecedem o gargalo mais uma folga para proteção). A Equação 7.7 mostra a fórmula para o cálculo do momento de liberação (*release date* – RD). Ao chegar um novo *job* no sistema, ele fica aguardando no *pre-shop pool* até o RD. Caso não haja estoque em processo (ordens aguardando processamento no gargalo), libera-se a ordem imediatamente.

$$RD = a_i + BLG - tbg \tag{7.7}$$

O RD foi considerado o mesmo para as ordens de fabricação dos componentes na linha crítica e não-crítica a fim de condicionar o sistema ao ritmo do gargalo. No cenário do gargalo localizado na máquina 1, não foi considerado o *tbg*, tendo em vista que não há restrição de suprimento. Após a entrada da ordem no chão de fábrica, seguiu-se a priorização FCFS nas filas das máquinas. Como todos os *jobs* possuem a mesma sequência de operações no roteiro de produção, não foi considerado o *Buffer Management*.

Os *outputs* do modelo DBR são: tempo médio de fluxo, *lead time* médio, utilização do gargalo, percentual de ordens atrasadas e atraso médio. A Figura 21 mostra uma representação do modelo conceitual do DBR.

Figura 21 - Modelo conceitual da simulação do sistema DBR



Fonte: Elaboração própria.

7.2 Modelo computacional DBR

Para desenvolvimento do modelo de simulação do DBR também foram empregados os elementos do *software* descritos no Apêndice A e utilizados os *inputs* apresentados na seção 5.3. O modelo computacional do DBR é bem parecido com o do CONWIP, como pode-se observar a representação do sistema de montagem na captura de tela apresentada na Figura 22. Contudo, foram adotados diferentes comandos, além de criadas outras variáveis e parâmetros. As principais diferenças entre os modelos estão entre a entrada e a liberação dos *jobs*.

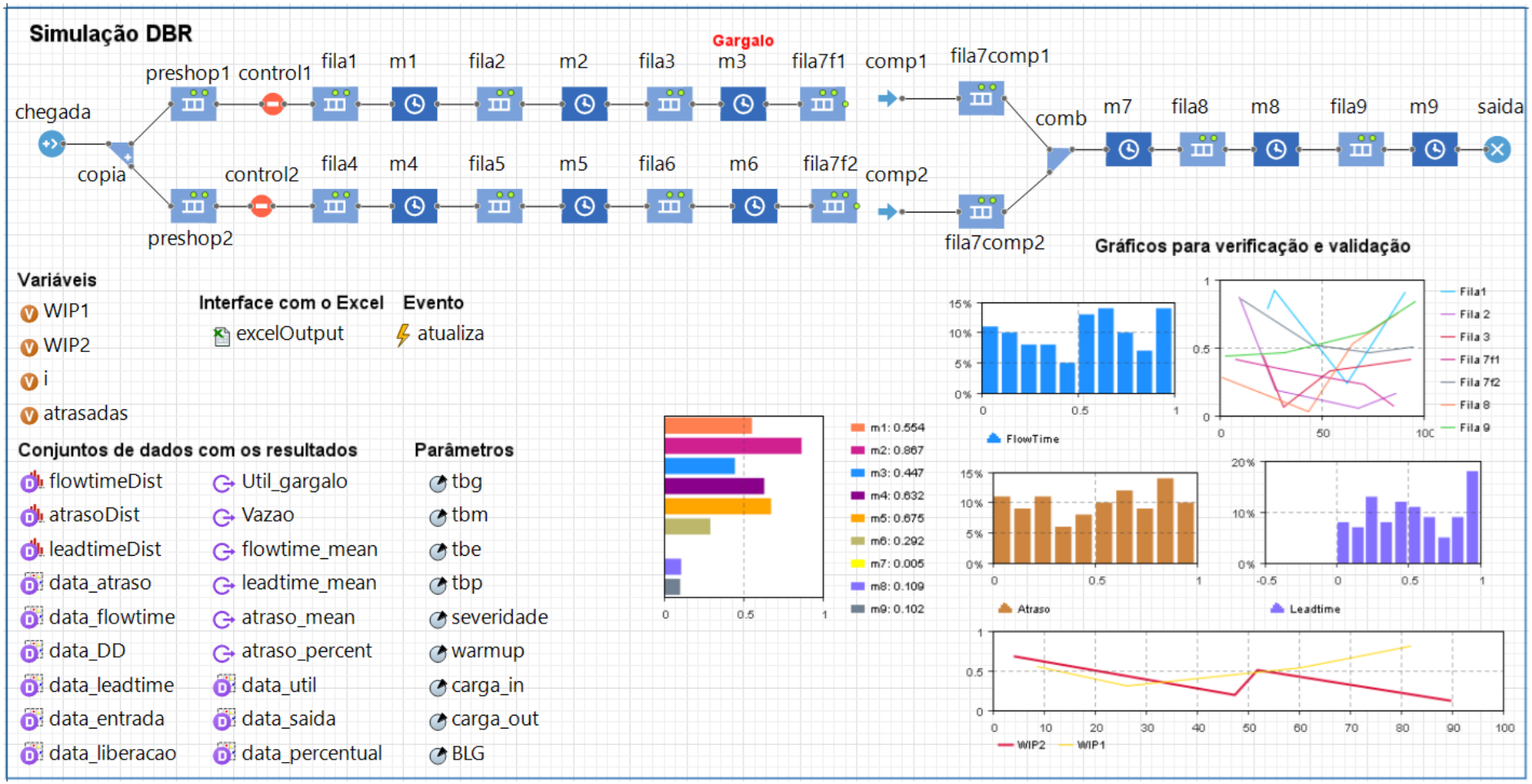
No bloco "*chegada*", são gerados os *jobs* e criados os atributos (identificação da ordem, data prometida e tempos médios de atendimento), como descrito no modelo CONWIP. Nesse momento, calcula-se também o PRD' e, caso o gargalo esteja localizado na máquina 1 ou 3, o PRD". Em seguida, o *job* passa pelo o bloco "*copia*", que tem a função de duplicar a entidade e transcrever todos os atributos da entidade original na cópia. Esse artifício foi usado para criação das ordens de fabricação.

As ordens de fabricação seguem para o *pre-shop pool* de cada linha (*preshop1* e *preshop1*), onde são ordenadas conforme menor momento planejado para liberação (PRD' e PRD", se houver). Nesse elemento, elas recebem o RD, conforme o cálculo da carga de trabalho já liberada para o gargalo (BLG). Configurou-se os blocos "*control1*" e "*control2*" para controlar a entrada, verificando continuamente se o RD da ordem era atingido. Para esse monitoramento contínuo, foi criado um evento chamado "*atualiza*", que a cada fração 0,001 unidades de tempo, verificava as condições do controle.

Após a liberação e processamento nas linhas de fabricação (m1-m3 e m4-m6), as ordens seguem para a montagem. De forma semelhante ao exposto no CONWIP, utilizou-se os elementos "*comp1*" e "*comp2*", juntamente com comandos lógicos nas filas "*fila7f1*" e "*fila7f2*" para garantir que os componentes de uma mesma ordem fossem unidos no processo de montagem.

Na Figura 22, abaixo do processo de montagem também estão algumas variáveis, parâmetros (como o *tbm – time buffer* da montagem expandido, *tbg – time buffer* do gargalo e *tbe – time buffer* da expedição), além de listas de dados, indicadores e gráficos utilizados para coleta e análise dos resultados da simulação. Esses elementos, assim como no CONWIP, auxiliaram o processo de verificação e validação do modelo.

Figura 22 - Captura da tela principal do modelo de simulação do DBR



Fonte: Elaboração própria.

7.3 Verificação e validação DBR

As estratégias descritas na seção 6.3 para verificação e validação do modelo CONWIP, sugeridas por Sargent (2013), Stewart (2004) e Law e Kelton (2000), também foram aplicadas na modelagem do DBR. São essas: (a) animações; (b) gráficos de desempenho; (c) relacionamento entre os dados; (d) testes degenerativos; (e) *trace*; (f) comparação dos resultados da simulação com os resultados esperados.

Além disso, foram realizados testes na simulação do DBR com taxa de chegada e tempos de processamento constantes. Isso foi importante para averiguar se o modelo realizava adequadamente o controle da liberação das ordens e comparar com resultados esperados.

A Tabela 2 mostra os resultados de um exemplo com o gargalo na máquina 3. Considerou-se o intervalo entre chegadas igual a 0,8 unidades, o tempo médio de processamento nas máquinas não críticas equivalente a 0,5 unidades e no gargalo (p_g) igual a 1,0 unidade de tempo. Isso leva a uma taxa de ocupação superior a 100%, no entanto, serão discutidos apenas os registros dos primeiros 5 *jobs* atendidos no sistema. A determinação da data prometida foi aleatória, como previsto na seção 5.3. O *time buffer* da montagem expandido foi de 3 unidades de tempo, enquanto do gargalo e o de expedição foram de 1,5 unidades cada.

Os PRDs foram calculados de acordo com as Equações 7.3 e 7.4. Por exemplo, para a ordem 1, que recebeu 29,01 como a data prometida, deduzindo desse valor o *tbe* = 1,5, o tempo de processamento no gargalo = 1,0 e o *tbg* = 1,5, tem-se 24,71 como valor do PRD'. Para o PRD", foi deduzido da data prometida apenas o *tbm* = 3,0, logo tem-se 25,01.

Para definição do momento de liberação (RD), aplicou-se a Equação 7.7, que soma o *backlog* do gargalo ao momento de chegada do *job* e subtrai o *tbg*. No caso da ordem 1, como não havia BLG, o modelo libera imediatamente para o chão de fábrica. Para o cálculo do RD da segunda ordem, que chegou no momento 1,6, foi adicionado o tempo de processamento da ordem anteriormente liberada (1,0) e subtraiu o *tbg* (1,5). Nesse caso, como o valor do RD é inferior ao momento de chegada, a ordem também foi liberada prontamente.

Observa-se que, a ordem 6 chegou no instante 4,8 e haviam duas ordens entre a fila da máquina 1 e o gargalo (ordens 4 e 5, equivalente ao BLG=2). Logo, o RD foi o programado para o momento 5,3 e a ordem ficou aguardando no *pre-shop pool*. As ordens 7, 8, 9 e 10 também esperaram na fila de entrada, a fim de reduzir o congestionamento da linha.

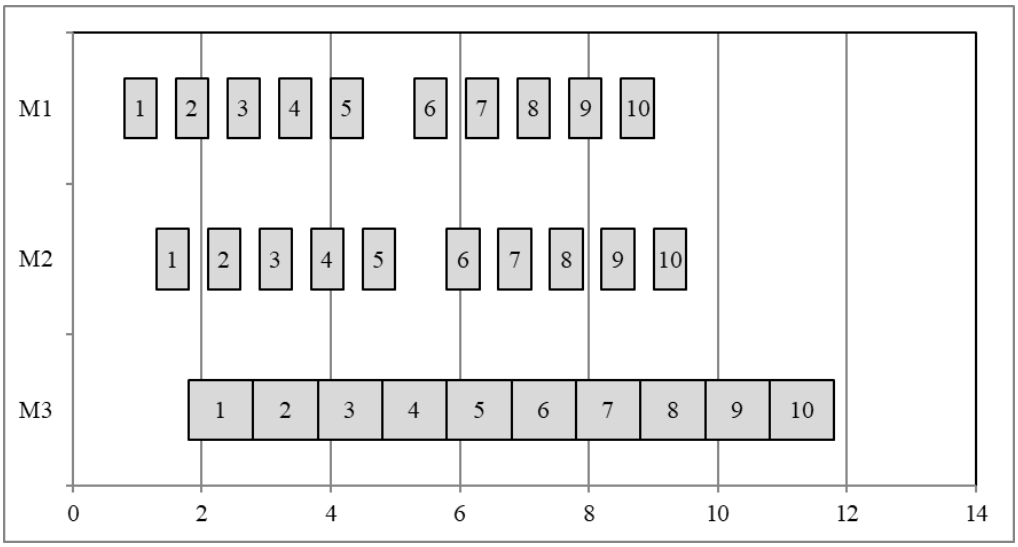
Tabela 2 - Exemplo de teste da verificação do DBR

OF	DD	PRD'	PRD''	RD	Entrada	Liberação	*tbg*	*tbm*	*tbe*	p_g
1	29,01	25,01	26,01	0,00	0,80	0,80	1,5	3	1,5	1
2	26,70	22,70	23,70	1,10	1,60	1,60	1,5	3	1,5	1
3	28,48	24,48	25,48	1,90	2,40	2,40	1,5	3	1,5	1
4	27,44	23,44	24,44	2,70	3,20	3,20	1,5	3	1,5	1
5	29,50	25,50	26,50	3,50	4,00	4,00	1,5	3	1,5	1
6	33,58	29,58	30,58	5,30	4,80	5,30	1,5	3	1,5	1
7	33,85	29,85	30,85	6,10	5,60	6,10	1,5	3	1,5	1
8	33,20	29,20	30,20	6,90	6,40	6,90	1,5	3	1,5	1
9	36,02	32,02	33,02	7,70	7,20	7,70	1,5	3	1,5	1
10	32,49	28,49	29,49	8,50	8,00	8,50	1,5	3	1,5	1

Fonte: Elaboração própria.

A Figura 23 apresenta o gráfico de Gantt da programação das 10 ordens até a máquina gargalo (m3). Observa-se que todas as ordens chegaram com antecedência na fila do recurso crítico, de modo que esse recurso não ficou aguardando para o processamento.

Figura 23 - Gráfico de Gantt da programação das ordens até o gargalo (M3)



Fonte: Elaboração própria.

CAPÍTULO 8 – DELINEAMENTO DOS EXPERIMENTOS

Este capítulo apresenta o delineamento dos experimentos desta pesquisa. Inicialmente, detalham-se os fatores e indicadores (variáveis de resposta) utilizados nos experimentos. Em seguida, descreve-se o processo de determinação da duração das rodadas de simulação e tempo de aquecimento, por meio da análise dos regimes transitório e permanente. Discute-se ainda a quantidade de replicações dos cenários com a análise do intervalo de confiança para as estimativas das variáveis. Por fim, são definidos os níveis do limite de estoque (um dos fatores dos experimentos) para os modelos CONWIP e DBR.

8.1 Fatores e variáveis de resposta

O delineamento dos experimentos objetiva permitir a comparação dos sistemas de programação e controle CONWIP e DBR em um *assembly flow shop*. Nesse contexto, os fatores do experimento são:

a) Método de programação e controle da liberação de ordens (CONWIP e DBR), foco da análise comparativa deste estudo.
b) Posição do gargalo no sistema: estudos anteriores sugerem que há diferenças no desempenho do CONWIP e DBR de acordo com a posição da estação gargalo em *flow shops* (KIM; DAVIS; COX III, 2003; LEE; SEO, 2016; THÜRER et al., 2017b), assim como, a localização do gargalo afeta o desempenho de *assembly shops* (FRY et al., 1987; DUENYAS, 1994a).
c) Limite de *buffer* (maxWIP ou *time buffer*): o limite de WIP no sistema afeta o desempenho dos modelos de controle em relação ao tempo de fluxo e produtividade (KIM; COX; MABIN, 2010; THÜRER; STEVENSON, 2018).
d) Severidade do gargalo: a severidade está relacionada com a estabilidade do sistema e, consequentemente, a possibilidade de flutuação do gargalo, além disso é um fator que pode comprometer o desempenho do ambiente (BAHOUTH; FOOTE, 1994; THÜRER; STEVENSON, 2018).

Os níveis dos fatores posição do gargalo e severidade do gargalo, já foram discutidos no Capítulo 5 (seção 5.3) durante a caracterização do ambiente em estudo. A posição possui quatro níveis: o gargalo pode estar localizado nas máquinas 1, 3, 7 ou 9. Enquanto que o fator

severidade possui três níveis: severidade alta (50%), média (30%) ou baixa (10%). O fator limite de *buffer* possui dez níveis. Os valores foram definidos após uma experimentação inicial, a fim de identificar a faixa em que os modelos apresentaram melhor desempenho, como será detalhado nas seções 8.3 e 8.4.

Para avaliação do desempenho do sistema de montagem em estudo, os seguintes indicadores foram considerados:

a) Utilização do gargalo: calculado como a razão entre o tempo em que o recurso crítico esteve ocupado em relação a disponibilidade do mesmo;
b) Tempo médio de fluxo (*flow time* médio): dado pela média das diferenças entre o momento de conclusão das ordens na máquina 9 (última operação) e o momento da liberação do primeiro componente das ordens para uma das linhas de fabricação;
c) *Lead time* médio (tempo médio de atravessamento): calculado pela média das diferenças entre o momento de conclusão das ordens na última máquina e o instante de chegada do *job* no sistema;
d) Percentual de ordens atrasadas: obtido pela razão da contagem das ordens finalizadas após a data prometida em relação ao quantitativo de ordens processadas no sistema;
e) Atraso médio: dado pela soma dos valores de atraso dos *jobs* divido pela quantidade total de *jobs* produzidos.

A utilização foi escolhida como um dos indicadores para os experimentos, pois como exposto em Chakravorty e Atwater (2005) e Atwater e Chakravorty (2002), a literatura de TOC enfatiza um alto valor desse indicador para o máximo desempenho do gargalo, apesar da utilização 100% não ser obtida na prática. Essa medida é comumente aplicada para avaliação do desempenho em estudos sobre o DBR e em estudos comparativos dos modelos CONWIP e DBR (GILLAND, 2002; KOH; BULFIN, 2004; MILLSTEIN; MARTINICH, 2014).

O tempo médio de fluxo também é um indicador muito utilizado em estudos sobre controle de liberação de ordens, especialmente, quando se deseja avaliar o desempenho quanto ao nível de estoque em processo (CHAKRAVORTY; ATWATER, 2005; SIRIKRAI; YENRADEE, 2006; THÜRER et al., 2017c; THÜRER; STEVENSON, 2018).

O *lead time* complementa a análise do tempo de fluxo, indicando o tempo total de atravessamento, que inclui o tempo que a ordem ficou aguardando pela liberação no *pre-shop pool*. Essa medida de desempenho também é frequentemente aplicada para avaliação dos sistemas em questão (THÜRER et al., 2017b; THÜRER; STEVENSON, 2018).

O percentual de atraso e atraso médio foram considerados em virtude da avaliação da confiabilidade de entrega dos modelos. Essas medidas são comumente avaliadas no contexto

MTO, tendo em vista que o desempenho de entrega é o fator crítico de sucesso dessa estratégia. Além disso, outros estudos já compararam o desempenho dos modelos CONWIP e DBR em relação a essas medidas (FREDENDALL; OJHA; PATTERSON, 2010; THÜRER et al., 2017b).

A combinação dos fatores dos experimentos resultou em 240 instâncias do problema (2x4x10x3) a serem simuladas. O próximo tópico detalha o processo de determinação de três parâmetros importantes para a experimentação: a duração dessas simulações, o tempo de aquecimento e a quantidade de replicações de cada cenário.

8.2 Determinação da duração das simulações, tempo de aquecimento e quantidade de replicações

Para determinar a duração das simulações e o tempo de aquecimento, utilizou-se a estratégia de rodar o modelo por um período de tempo longo (30.000 unidades de tempo) a fim de observar o regime transitório e permanente das variáveis de resposta. Esse período permitiu a conclusão de aproximadamente 27 mil *jobs*, coerente com a taxa de chegada no sistema de 0,9 *jobs*/unidade de tempo.

Para efeito do dimensionamento dos experimentos, foram analisadas duas instâncias equivalentes, uma do CONWIP e outra do DBR, com os seguintes parâmetros: gargalo localizado na máquina 3, severidade média, tempos médios de atendimentos e tolerância para data prometida como já discutidos na seção 5.3. Quanto ao limite de WIP, no CONWIP adotou-se 15 *jobs* nos dois segmentos de maxWIP. No DBR, considerou-se uma proteção semelhante, equivalente a 15 unidades de tempo.

A cada conclusão de um pedido foram registrados os valores médios dos indicadores quantitativos (tempo de fluxo, *lead time* e atraso) e o percentual de ordens atrasadas. Os gráficos com os valores médios das variáveis ao longo do tempo de simulação dos dois modelos são apresentados no Apêndice B.

Após observar o comportamento das cinco variáveis (análise dos regimes transitório e estacionário), decidiu-se considerar 10.000 unidades de tempo de simulação, sendo o *warm-up* de 3.000 unidades de tempo. Esses tempos de aquecimento e de duração das simulações são semelhantes aos adotados em outros estudos que avaliaram ambientes configurados de forma semelhante ao desta pesquisa (THÜRER et al., 2012, 2017c; THÜRER; STEVENSON, 2018).

Após a definição do *warm-up* e do tempo de duração das simulações, as mesmas instancias foram simuladas por diversas vezes com diferentes sementes a fim de analisar a confiabilidade dos resultados. A Tabela 3 apresenta os resultados finais dos indicadores para uma série de 50 replicações (as primeiras e as últimas dez linhas) para o experimento com o CONWIP, enquanto que a Tabela 4 mostra os resultados para o sistema DBR. Foram calculados os desvios, as médias e as precisões para determinação dos intervalos de confiança com nível de 95%. Os valores de precisão foram considerados aceitáveis, de modo que, essa quantidade de replicações foi adotada nos experimentos.

Tabela 3 - Cálculo dos intervalos de confiança com base em 50 replicações do modelo CONWIP

Replicação	Atraso médio	Percentual de atraso	*Flow time* médio	*Lead time* médio	Utilização	Posição	maxWIP	Severidade
1	0,390	4,617%	11,232	13,366	0,908	3	15	Média
2	2,034	12,250%	11,553	16,474	0,928	3	15	Média
3	0,003	0,234%	10,063	10,686	0,893	3	15	Média
4	0,320	4,130%	10,651	12,345	0,891	3	15	Média
5	0,015	0,535%	10,148	11,035	0,896	3	15	Média
6	0,002	0,170%	10,236	10,793	0,881	3	15	Média
7	0,005	0,166%	10,683	11,418	0,909	3	15	Média
8	0,021	0,516%	10,113	10,741	0,890	3	15	Média
9	0,290	2,454%	10,071	11,303	0,880	3	15	Média
10	0,040	0,788%	10,155	10,836	0,899	3	15	Média
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
40	0,129	2,283%	10,504	11,796	0,906	3	15	Média
41	0,006	0,280%	10,456	11,171	0,896	3	15	Média
42	0,015	0,557%	10,513	11,365	0,902	3	15	Média
43	0,083	1,768%	10,879	12,167	0,896	3	15	Média
44	0,795	4,309%	10,701	12,951	0,901	3	15	Média
45	0,009	0,538%	10,250	11,075	0,894	3	15	Média
46	0,018	0,810%	10,965	12,089	0,908	3	15	Média
47	0,075	1,547%	10,040	10,809	0,886	3	15	Média
48	0,188	2,176%	10,467	11,606	0,898	3	15	Média
49	0,026	0,925%	10,127	10,910	0,890	3	15	Média
50	0,213	2,404%	10,899	12,601	0,894	3	15	Média
Desvio padrão	0,344	2,336%	0,456	1,188	0,011			
Média	0,212	2,354%	10,578	11,944	0,900			
Precisão	**0,098**	**0,664%**	**0,130**	**0,338**	**0,003**			
Intervalo de confiança (95%)	**0,114≤ μ ≤0,310**	**1,69%≤ *p* ≤3,018%**	**10,448≤ μ ≤10,707**	**11,606≤ μ ≤12,282**	**0,897≤ μ ≤0,903**			

Fonte: Elaboração própria.

Tabela 4 - Cálculo dos intervalos de confiança com base em 50 replicações do modelo DBR

Replicação	Atraso médio	Percentual de atraso	*Flow time* médio	*Lead time* médio	Utilização	Posição	*Time buffer*	Severidade
1	0,017	0,48%	10,150	10,646	0,876	3	15	Média
2	0,146	2,47%	10,090	10,686	0,894	3	15	Média
3	0,006	0,22%	10,347	10,836	0,897	3	15	Média
4	0,005	0,25%	10,539	11,023	0,893	3	15	Média
5	0,039	1,00%	10,559	11,149	0,904	3	15	Média
6	0,059	1,46%	10,882	11,518	0,905	3	15	Média
7	0,013	0,56%	10,744	11,306	0,896	3	15	Média
8	0,658	4,85%	10,941	12,239	0,907	3	15	Média
9	0,035	0,98%	10,617	11,284	0,908	3	15	Média
10	0,261	2,80%	11,017	11,895	0,902	3	15	Média
(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
40	0,068	1,44%	10,448	11,092	0,905	3	15	Média
41	0,975	6,85%	12,573	14,488	0,910	3	15	Média
42	0,009	0,44%	10,442	10,846	0,901	3	15	Média
43	0,012	0,48%	9,941	10,328	0,905	3	15	Média
44	0,043	1,27%	10,851	11,627	0,899	3	15	Média
45	0,030	0,99%	10,833	11,419	0,911	3	15	Média
46	0,468	6,07%	12,193	13,738	0,910	3	15	Média
47	0,002	0,20%	10,255	10,687	0,877	3	15	Média
48	1,208	8,12%	13,126	15,344	0,910	3	15	Média
49	0,063	1,53%	10,339	10,919	0,896	3	15	Média
50	0,067	1,86%	10,878	11,657	0,901	3	15	Média
Desvio Padrão	0,251	1,945%	0,730	1,115	0,012			
Média	0,165	2,083%	10,930	11,758	0,901			
Precisão	**0,071**	**0,553%**	**0,207**	**0,317**	**0,003**			
Intervalo de confiança (95%)	**0,094≤ μ ≤0,236**	**1,53%≤ *p* ≤2,636%**	**10,722≤ μ ≤11,137**	**11,441≤ μ ≤12,075**	**0,898≤ μ ≤0,905**			

Fonte: Elaboração própria.

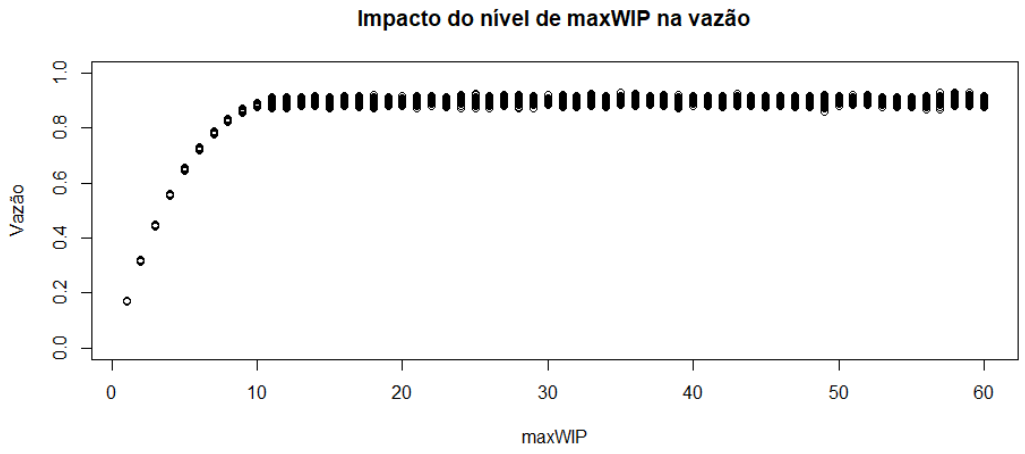
8.3 Determinação dos níveis de maxWIP para o modelo CONWIP

Para concluir a etapa de delineamento dos experimentos, foi preciso avaliar os limites de *buffer* a serem considerados nos modelos CONWIP e DBR. Diferentemente dos outros fatores, os níveis de *buffer* foram definidos após uma experimentação inicial dos modelos a fim de identificar faixas de valores de maxWIP e *time buffer* que resultassem em um bom desempenho para os modelos.

No caso do CONWIP, considerou-se, inicialmente, que os dois segmentos possuiriam um mesmo nível de WIP (maxWIP1=maxWIP2, que nessas condições será chamado apenas maxWIP neste texto). Esse foi o único fator que variou no experimento. Foram simulados cenários com maxWIP de 1 a 60 unidades com duração de 10.000 unidades de tempo e *warm-up* de 3.000 unidades de tempo, sendo replicados 50 vezes (como definido na seção anterior). O gargalo era a terceira máquina do sistema de montagem com severidade baixa (10%).

O efeito da variação do nível do maxWIP foi observado quanto a vazão do sistema, tempo de fluxo médio e *lead time* médio. Observa-se na Figura 24, que a vazão do sistema é crescente nos níveis iniciais, mas algumas unidades após o maxWIP=10, estabiliza-se em torno de 0,90 *jobs*/unidade de tempo, igual a taxa de chegada dos pedidos. Esse comportamento é semelhante ao observado na análise da utilização do gargalo com a taxa média de 90% de utilização.

Figura 24 - Análise inicial do impacto do nível de maxWIP na vazão do sistema

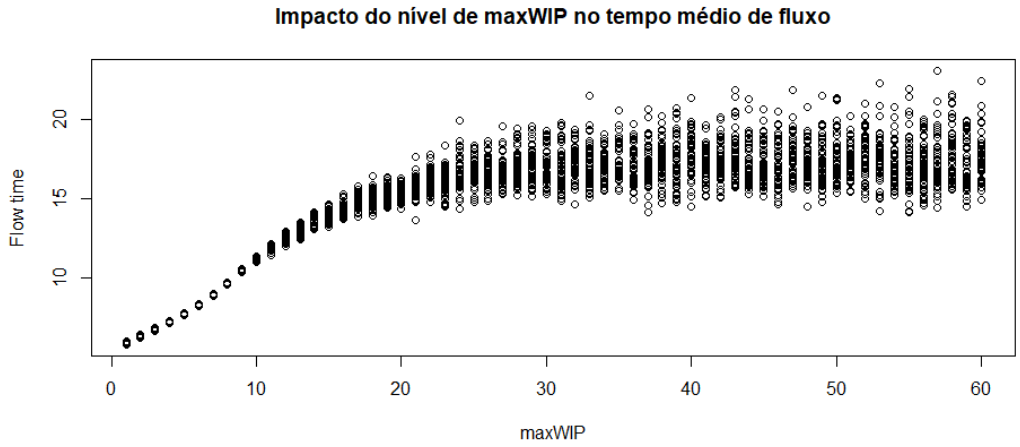


Fonte: Elaboração própria.

Em relação ao tempo de fluxo, verifica-se na Figura 25, que o tempo médio de permanência das ordens no chão de fábrica também foi crescente, especialmente nos primeiros níveis de WIP, e posteriormente, apresentou variações concentradas em torno de 15 e 20

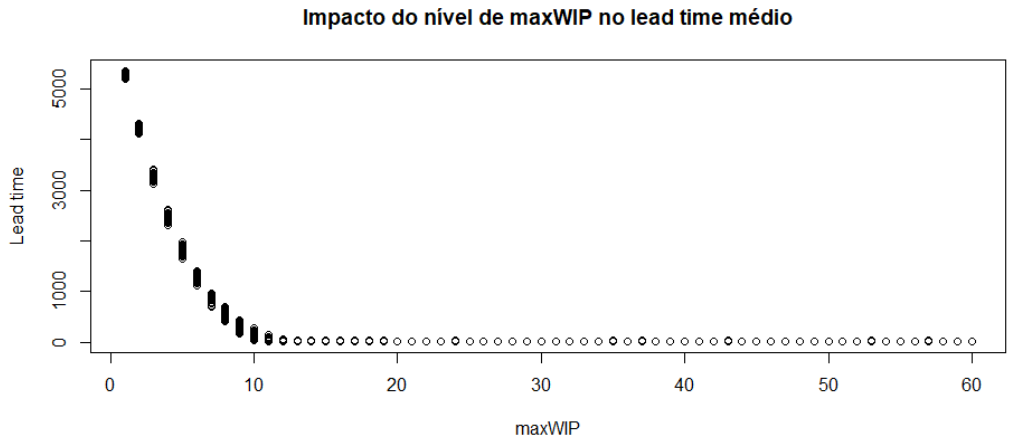
unidades de tempo. Enquanto que o *lead time* médio, como exposto na Figura 26, apresentou valores altíssimos, quando a restrição de WIP no sistema era mais severa. No entanto, um pouco depois do maxWIP=10, os valores médios se tornaram menores e mais estáveis, se aproximando do tempo de fluxo.

Figura 25 - Análise inicial do impacto do nível de maxWIP no tempo médio de fluxo do sistema



Fonte: Elaboração própria.

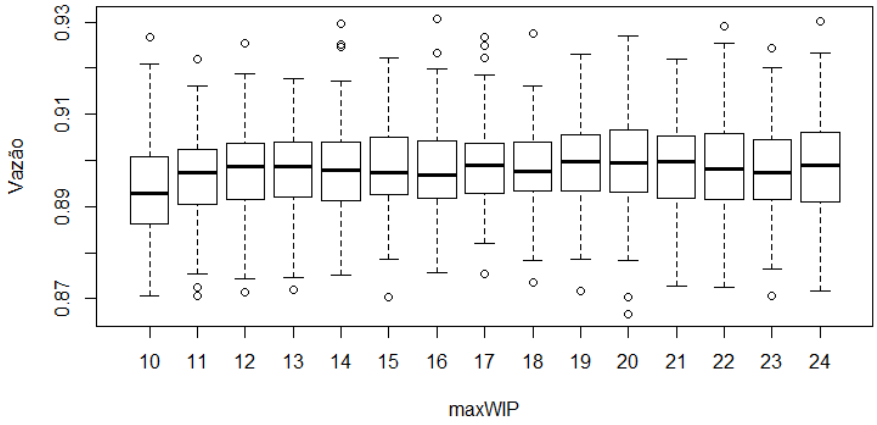
Figura 26 - Análise inicial do impacto do nível de maxWIP no *lead time* médio do sistema



Fonte: Elaboração própria.

Diante dessa análise inicial, decidiu-se investigar o impacto da variação do maxWIP entre 10 e 24 unidades considerando todos os níveis de severidade do gargalo (10%, 30% e 50%). A Figura 27 apresenta o gráfico *boxplot* da vazão do sistema. Pode-se observar que esse indicador apresenta variação entre 0,87 e 0,93 com a média em torno de 0,9 (o que remete a uma taxa de 90% de ocupação do gargalo). Isso demostra que, para esse indicador, o sistema está oscilando próximo a condição de melhor desempenho.

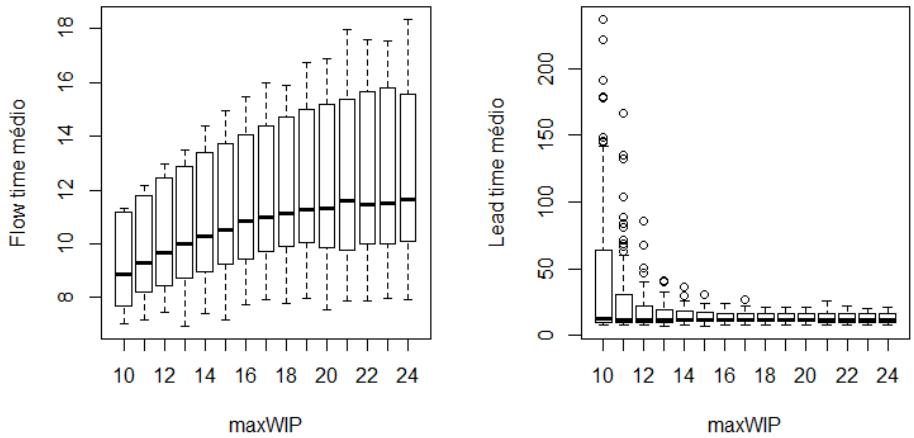
Figura 27 - Análise inicial do efeito do nível do maxWIP (10 a 24) na vazão do sistema



Fonte: Elaboração própria.

Em relação aos indicadores tempo médio de fluxo e tempo médio de atravessamento, verifica-se na Figura 28 que nessa faixa de maxWIP há um certo grau de variação dos resultados. O *flow time* apresentou valores entre 7,5 e 18 unidades de tempo com a média das replicações em torno de 10 e 11 unidades de tempo. Nos níveis iniciais de maxWIP, o *lead time* apresentou valores muito altos, o que comprometeu também o desempenho das ordens em termos de atraso médio e percentual de atraso. O *lead time* começou a se estabilizar após o limite de WIP igual a 12 ordens, ao passo que o tempo de fluxo seguiu crescente.

Figura 28 - Análise inicial do efeito do nível do maxWIP (10 a 24) no *flow time* e *lead time* médios



Fonte: Elaboração própria.

Essa mesma análise inicial foi realizada com outras posições do gargalo: a primeira e a última máquina do sistema. Como os resultados apresentaram desempenho semelhante, definiu-se que os dez níveis considerados nos experimentos de maxWIP seriam de 12 a 21 ordens.

Após essa definição dos limites de WIP, avaliou-se a relação entre o maxWIP1 e maxWIP2. Foram realizados testes com as constantes: 0,8, 0,9, 1,0, 1,1 e 1,2 multiplicados pelo maxWIP1 para cálculo do parâmetro maxWIP2 (maxWIP2 = constante x maxWIP1). Considerou-se, inicialmente, o caso do gargalo localizado na máquina 3 em todos os níveis de severidade.

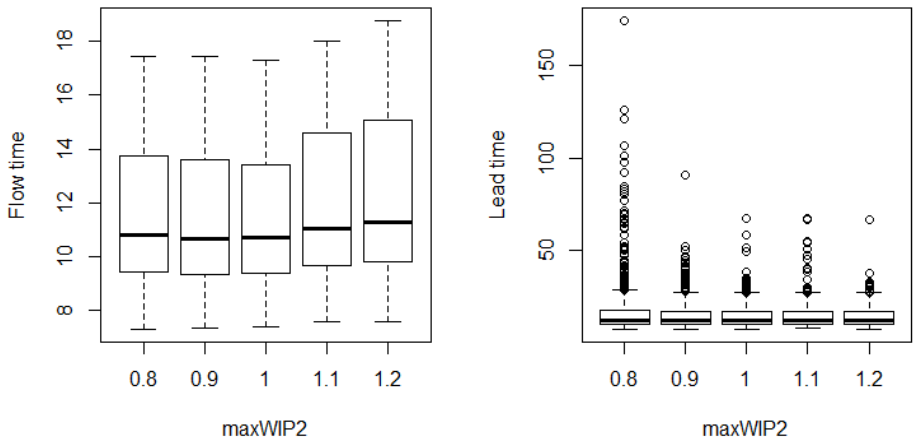
Quando a constante era igual 1,0 significava que os limites eram iguais e as ordens eram liberadas simultaneamente nas duas linhas, como testado anteriormente. Quando menor que 1,0 (maxWIP2<maxWIP1), as ordens eram liberadas primeiro na linha crítica (com o gargalo) e depois na linha não-crítica. Caso maior que 1,0 (maxWIP2>maxWIP1), liberava primeiro as ordens na linha não-crítica e depois na linha de fabricação com o gargalo.

Como esperado, os resultados indicaram que diferentes níveis de WIP no segmento 2 não afetam a utilização do gargalo, tendo em vista que esse está localizado no segmento 1. As principais diferenças foram observadas quanto ao tempo médio de fluxo e ao *lead time* médio.

Observa-se na Figura 29 que para o sistema em estudo, diante das características das linhas de fabricação simuladas, não há um ganho expressivo no desempenho com a estratégia de "segurar um pouco mais" a liberação na linha livre (maxWIP2=0,8 ou 0,9) ou antecipá-la em relação à linha com o gargalo (maxWIP2=1,1 ou 1,2).

No primeiro caso, os tempos médios de fluxo foram parecidos, mas houve comprometimento do tempo de atravessamento das ordens. Com isso o percentual de atraso das ordens e o atraso médio também aumentaram. Enquanto isso, a estratégia de liberar antes as ordens na linha não-crítica não comprometeu o *lead time*, porém causou um aumento do tempo médio de fluxo das ordens.

Figura 29 - Análise da relação entre o maxWIP1 e o maxWIP2 quanto ao tempo médio de fluxo e o *lead time* médio



Fonte: Elaboração própria.

A análise da relação entre o maxWIP1 e maxWIP2 também foi feita para quando o gargalo está localizado na máquina 1 e os resultados foram muito semelhantes. Dessa forma, decidiu-se adotar um mesmo valor de maxWIP para os *loops*.

8.4 Determinação dos níveis de *time buffer* para o modelo DBR

Inicialmente, foram consideradas as referências clássicas para dimensionamento dos *time buffers* para as simulações do DBR. No entanto, como esta pesquisa é de natureza teórica, o ambiente simulado não possui *benchmarks* reais do *lead time*, tempo de fluxo e histórico de flutuações para avaliação como sugere a literatura.

Desse modo, o modelo CONWIP serviu de parâmetro para a determinação dos níveis de *time buffer* no DBR, a fim de que os dois modelos fossem executados em condições semelhantes e permitisse a comparação dos mesmos. Durante as simulações do CONWIP calculou-se o tempo gasto pelas ordens nos trechos do sistema que deveriam ser protegidos pelos *time buffers*.

Por exemplo, no caso em que a máquina 3 é o gargalo, devem ser considerados os três *buffers* (gargalo, expedição e montagem expandido). Logo, calculou-se o tempo médio gasto pelas ordens desde a liberação na linha crítica até o início do atendimento na operação 3 (estimativa para *time buffer* do gargalo). A diferença entre o momento de conclusão da ordem na máquina gargalo até a finalização na máquina 9 permitiu a estimativa para o *time buffer* da expedição. Por fim, o *time buffer* da montagem expandido foi calculado como a diferença entre o momento da liberação da ordem na linha não-crítica e a conclusão na máquina 9.

Essas medições serviram de *benchmark* para as estimativas dos níveis de *time buffer* no DBR. Os resultados de todos os cenários rodados do CONWIP foram analisados, segmentando-os de acordo com o maxWIP. A Tabela 5 apresenta as médias e os desvios das 50 replicações para os indicadores tempo médio de fluxo e *lead time* médio e para as estimativas de *time buffer* da montagem expandido (*tbm*), do gargalo (*tbg*) e da expedição (*tbe*), estratificados também pela posição do recurso crítico (*g*). Apenas o *time buffer* da montagem expandido não foi analisado por posição, pois o trecho protegido era o mesmo, independente se o recurso crítico era a máquina 1 ou 3.

A média do tempo médio de fluxo aferida com as 50 replicações ficou entre 10,24 e 12,39 unidades, enquanto que para o *lead time* a média oscilou entre 13,27 e 17,83 unidades. Observa-se na Tabela 5, que o *tbm* apresentou valores próximos ao tempo de fluxo, enquanto que o *tbg* e *tbe* representam frações do tempo de fluxo de acordo com a localização do gargalo. O *time*

buffer da expedição, por exemplo, é bem maior no caso quando o gargalo é a máquina 1 do que quando o gargalo é a máquina 7. Assim como, o *tbg* é menor no caso do gargalo na terceira estação em relação ao cenário com a restrição na última máquina do sistema.

Tabela 5 - Resumo do resultado de 50 replicações do CONWIP para *benchmark* dos *time buffers* no DBR

maxWIP	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**
***Flow time* (média)**	10,24	10,62	10,96	11,25	11,52	11,79	11,97	12,18	12,30	12,39
***Flow time* (desvio)**	1,78	1,90	2,03	2,17	2,30	2,45	2,54	2,65	2,70	2,78
***Lead time* (média)**	17,83	15,23	14,26	13,97	13,70	13,65	13,55	13,51	13,37	13,27
***Lead time* (desvio)**	11,89	6,87	5,13	4,56	4,21	4,05	3,91	3,86	3,61	3,54
***tbm* (média)**	10,12	10,50	10,84	11,11	11,36	11,58	11,78	12,02	12,14	12,25
***tbm* (desvio)**	1,80	1,91	2,03	2,18	2,36	2,50	2,58	2,66	2,72	2,74
***tbg*, g=3 (média)**	5,58	5,90	6,20	6,42	6,65	6,81	7,06	7,28	7,33	7,52
***tbg*, g= 3 (desvio)**	0,30	0,36	0,45	0,56	0,68	0,78	0,83	0,89	0,96	1,00
***tbg*, g=7 (média)**	7,11	7,45	7,80	8,04	8,31	8,57	8,80	8,79	9,07	9,15
***tbg*, g=7 (desvio)**	0,80	0,91	1,04	1,15	1,23	1,32	1,44	1,58	1,61	1,72
***tbg*, g=9 (média)**	9,39	9,78	10,10	10,43	10,74	11,12	11,19	11,55	11,50	11,57
***tbg*, g=9 (desvio)**	1,77	1,86	2,00	2,13	2,18	2,39	2,46	2,58	2,63	2,84
***tbe*, g=1 (média)**	5,89	5,97	6,03	6,08	6,14	6,18	6,20	6,24	6,25	6,22
***tbe*, g=1 (desvio)**	2,69	2,80	2,87	2,96	3,03	3,09	3,13	3,18	3,19	3,15
***tbe*, g=3 (média)**	3,55	3,61	3,64	3,66	3,69	3,72	3,76	3,74	3,75	3,75
***tbe*, g=3 (desvio)**	1,67	1,74	1,80	1,83	1,86	1,90	1,96	1,93	1,94	1,94
***tbe*, g=7 (média)**	2,21	2,25	2,28	2,29	2,31	2,33	2,34	2,34	2,35	2,34
***tbe*, g=7 (desvio)**	0,97	1,02	1,06	1,08	1,10	1,16	1,15	1,15	1,16	1,15

Fonte: Elaboração própria.

Para não haver duplicação de fatores nos casos em que dois ou mais pulmões são aplicados, buscou-se correlacionar os valores dos *time buffers*. Adotou-se nos experimentos um parâmetro geral, chamado *time buffer* da produção (*tbp*). Os valores dos demais pulmões (*tbm*, *tbg* e *tbe*) tornam-se frações do valor do *tbp*.

Inicialmente, para observar o efeito dos *time buffers* nos indicadores, testou-se 14 níveis de *tbp*: 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 ou 20 unidades de tempo. Esses valores compreendiam os valores mínimo e máximo, além das médias do tempo de fluxo e ainda uma folga, como recomenda a literatura do DBR. O *tbm*, *tbg* e *tbe* foram dimensionados, multiplicando os valores do experimento para o *tbp* pelas constantes apresentadas na Tabela 6. Essas constantes foram estimadas após analisar a razão entre as estimativas médias dos *time buffers* e a média do tempo fluxo nos resultados da Tabela 5, sendo feita uma estimativa conservadora.

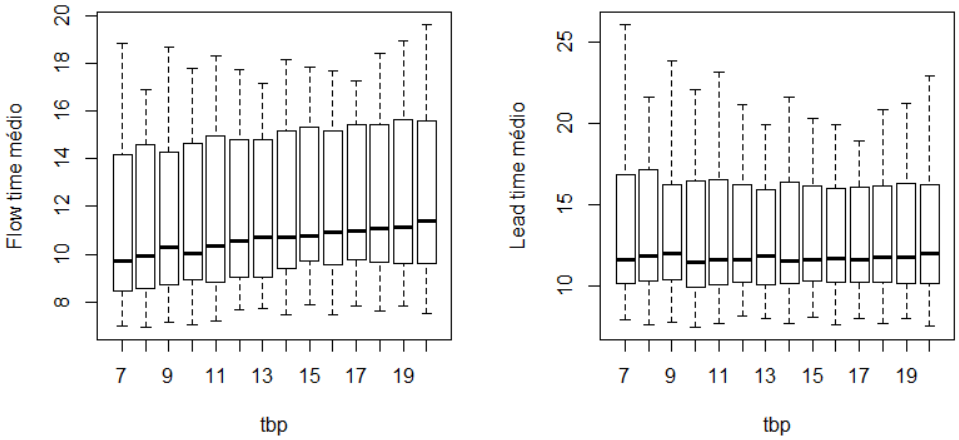
Tabela 6 - Constantes para dimensionamento dos *time buffers* em função da posição do gargalo

Posição	**Máquina1**		**Máquina 3**			**Máquina 7**		**Máquina 9**
Fatores	***tbm***	***tbe***	***tbm***	***tbg***	***tbe***	***tbg***	***tbe***	***tbg***
Constantes	1,00	1,00	1,00	0,65	0,35	0,85	0,30	1,10

Fonte: Elaboração própria.

A Figura 30 apresenta os resultados dos indicadores tempo médio de fluxo e *lead time* médio da simulação dos níveis de *tbp* com o gargalo localizado na máquina 3, ou seja, utilizando os três tipos de *time buffers* (*tbm*, *tbg* e *tbe*, conforme fatores da Tabela 6) em todas as severidades. Observa-se que o tempo médio de fluxo é crescente com o aumento do *tbp*, ou seja, quando se libera com maior antecedência as ordens. Enquanto que o *lead time*, teve maior dispersão dos resultados, mas com a mediana próxima a 12 unidades em todos os níveis.

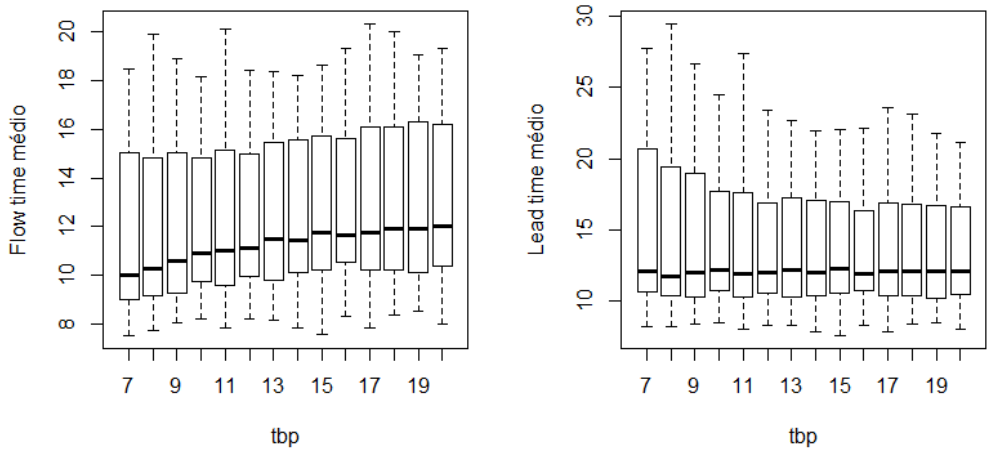
Figura 30 - Análise inicial do efeito do *time buffer* no *flow time* e *lead time* médios com o gargalo na máquina 3



Fonte: Elaboração própria.

A Figura 31 apresenta os resultados quando o gargalo é a máquina 9. O tempo médio de fluxo apresentou o mesmo comportamento quando o gargalo era a máquina 3 (crescente com o aumento da proteção). Para a variável *lead time* ficou evidenciou-se uma diminuição na dispersão dos resultados com o aumento do *tbp*.

Figura 31 - Análise inicial do efeito do *time buffer* no *flow time* e *lead time* médios com o gargalo na máquina 9



Fonte: Elaboração própria.

Nos dois casos, a utilização do recurso gargalo ficou em torno de 90% em todos os níveis de *tbp* analisados. Diante dessas constatações iniciais, foram feitas análises com intervalos menores de *tbp*. Então, determinou-se os dez níveis para avaliação nos experimentos: 9,5; 10; 10,5; 11; 11,5; 12; 12,5; 13; 13,5; 14. Esses valores, multiplicados pelas constantes da Tabela 6, são superiores as médias dos *time buffers* do gargalo identificados no CONWIP (Tabela 5), oferecendo proteção para eventuais variabilidades.

Com a definição dos valores para os limites de *buffer* aplicados nos modelos CONWIP e DBR, todos os fatores dos experimentos possuem o detalhamento dos níveis que serão analisados, como sintetiza a Tabela 7. Desse modo, o próximo Capítulo detalhará os resultados das simulações, conforme fatores, níveis e os parâmetros da duração das rodadas, tempo de aquecimento e quantidade de replicações.

Tabela 7 - Síntese dos fatores e níveis dos experimentos

Fatores	**Níveis**
Modelo de programação e controle da produção	CONWIP ou DBR
Limite de *buffer*	MaxWIP: 12, 13, 14, 15, 16, 17, 18, 19, 20, 21
	tbp: 9,5; 10; 10,5; 11; 11,5; 12; 12,5; 13; 13,5; 14
Severidade do gargalo	10%, 30% ou 50%
Posição do gargalo	Operação 1, 3, 7 ou 9

Fonte: Elaboração própria.

CAPÍTULO 9 – RESULTADOS E DISCUSSÕES

Este capítulo apresenta os resultados das simulações para as instâncias descritas no capítulo de Delineamento dos Experimentos. Inicialmente, é feita uma análise descritiva dos resultados dos modelos CONWIP e DBR. Posteriormente, os efeitos de cada um dos fatores no desempenho das variáveis de resposta são investigados com uma Análise de Variância (ANOVA – *Analysis of Variance*). Os resultados dos modelos são comparados e, por fim, são discutidos frente a literatura e a problematização do trabalho.

9.1 Execução da simulação

Os 240 problemas discutidos no delineamento dos experimentos foram replicados por 50 vezes, resultando em um total de 12.000 simulações. Cada uma com duração de 10.000 unidades de tempo e *warm-up* de 3.000 unidades de tempo. Os modelos foram executados em um computador com processador *Intel Core i7-7500U* e 16GB de memória RAM. Essa combinação de cenários levou aproximadamente 5 horas e meia para rodar no *AnyLogic* versão *Professional 8.3.3*.

9.2 Resultados do CONWIP

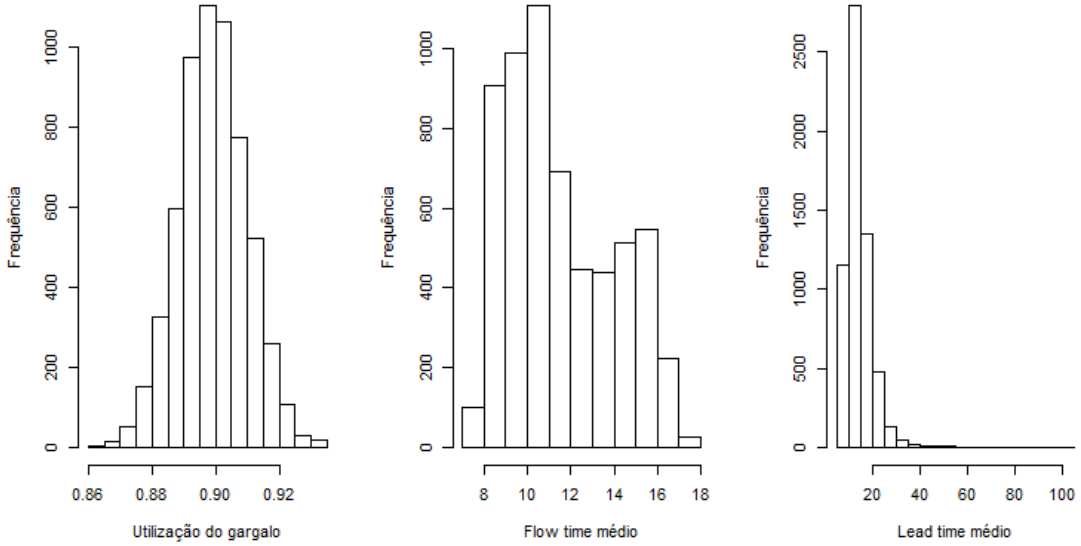
Com base nos resultados das simulações do CONWIP foram construídos histogramas das distribuições dos indicadores de desempenho: utilização do gargalo, tempo médio de fluxo, *lead time* médio, atraso médio e percentual de ordens atrasadas, como mostram as Figuras 32 e 33.

Quanto a utilização do gargalo, observa-se uma distribuição simétrica em torno de 90% da taxa de ocupação do recurso. Em relação ao tempo médio de fluxo, verifica-se que não há simetria nos resultados, havendo maior concentração de casos com valores entre 8 e 11 unidades de tempo. A distribuição do *lead time* médio é assimétrica positiva, havendo grande concentração de resultados inferiores a 20 unidades de tempo (Figura 32).

Quanto aos indicadores de confiabilidade de entrega, observa-se que ambas as distribuições são assimétricas positivas. O atraso médio apresentou maior frequência de

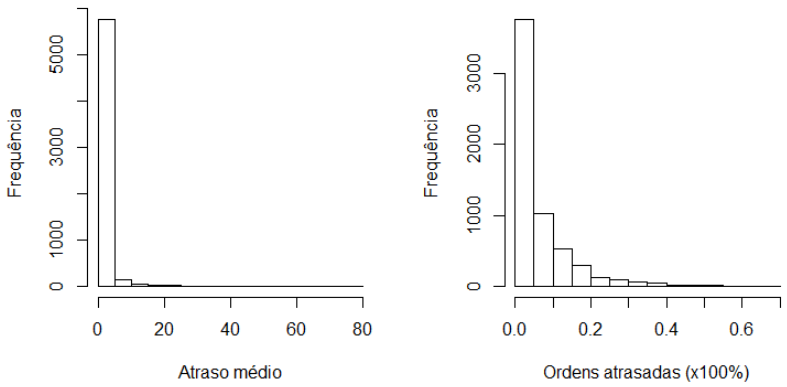
resultados inferiores a 5 unidade de tempo e o percentual de ordens atrasadas na maioria das simulações foi inferior a 10% (Figura 33).

Figura 32 - Distribuição dos resultados dos indicadores utilização do gargalo, tempo médio de fluxo e *lead time* médio no CONWIP



Fonte: Elaboração própria.

Figura 33 - Distribuição dos resultados dos indicadores atraso médio e percentual de ordens atrasadas no CONWIP



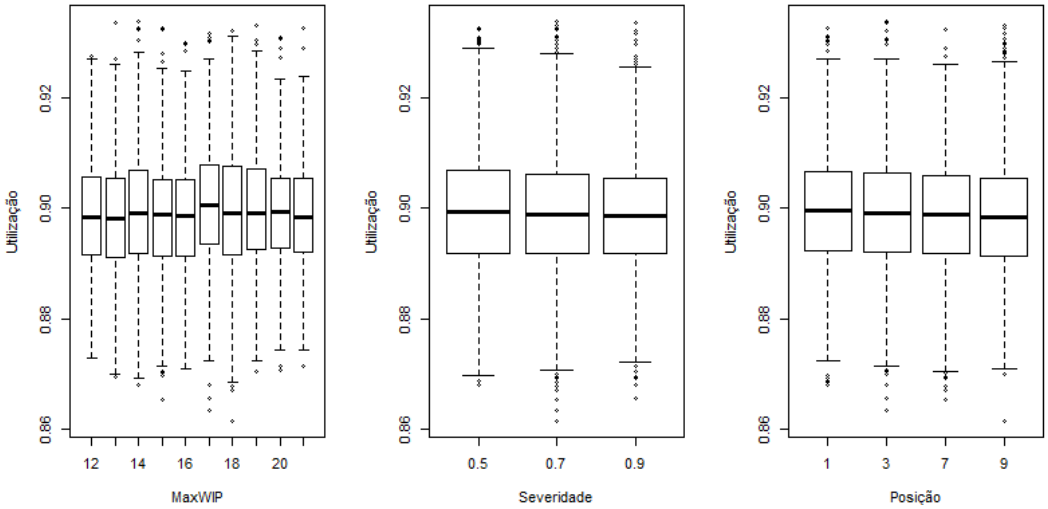
Fonte: Elaboração própria.

Em relação a variável utilização média do gargalo, a Figura 34 mostra que os fatores não influenciaram o nível de utilização do gargalo nas simulações do CONWIP. Em todos os níveis de maxWIP, a média ficou em torno de 90% com flutuações entre valores menores e maiores. O mesmo foi observado diante da severidade, com dispersão ligeiramente menor na severidade mais baixa (0,9) e na posição, com dispersão um pouco maior quando o gargalo era a máquina 3.

Em todos os casos, as diferenças foram mínimas. Isso indica que a utilização do gargalo foi determinada pela taxa de chegada das ordens e não pelos fatores. Por essa razão, os

resultados dos experimentos com o CONWIP foram estratificados pelos fatores em relação aos demais indicadores: tempo médio de fluxo, *lead time* médio, atraso médio e percentual de ordens atrasadas.

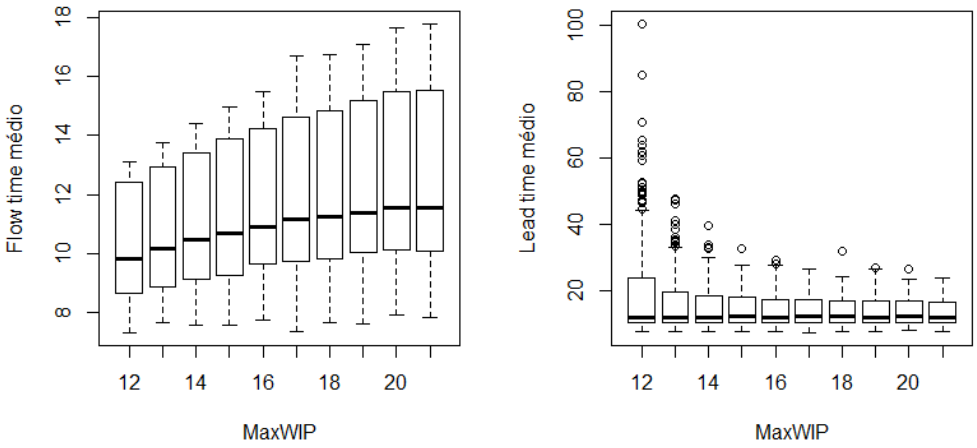
Figura 34 - Efeito dos fatores na utilização do gargalo no modelo CONWIP



Fonte: Elaboração própria.

Desse modo, pode-se analisar o efeito dos níveis de maxWIP nos indicadores tempo médio de fluxo e *lead time* médio (Figura 35) e no percentual de atraso e atraso médio (Figura 36). Como sugere a literatura (CONWAY et al., 1988; HOPP; SPEARMAN, 2013), há evidências de que um aumento do limite de WIP no sistema, eleva o tempo médio de fluxo (Figura 35). Em relação ao *lead time* médio, a mediana dos resultados foi semelhante em todos os níveis. No entanto, à medida que o maxWIP aumentou, a dispersão dos resultados diminuiu.

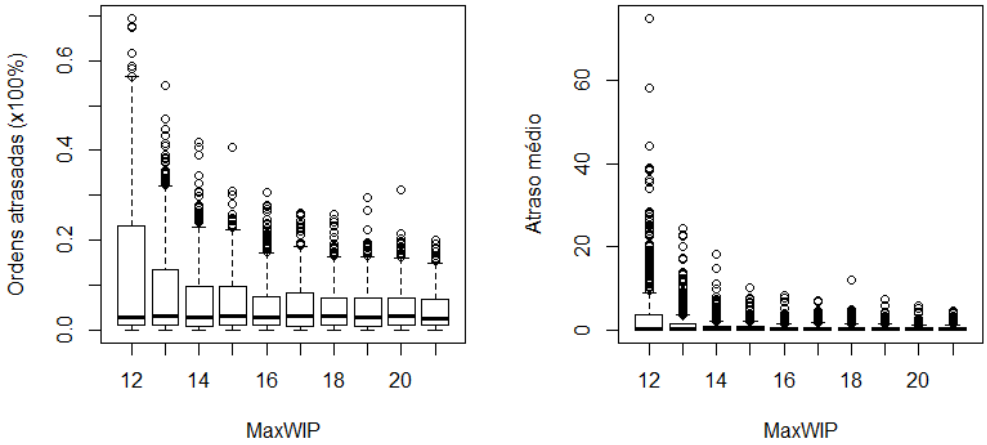
Figura 35 - Efeito dos níveis de maxWIP no *flow time* e *lead time* médios no modelo CONWIP



Fonte: Elaboração própria.

Observa-se na Figura 36, que a liberação de uma quantidade maior de *jobs* no sistema contribui para uma melhoria no desempenho de entrega. Quando o sistema operou com níveis menores de maxWIP, houve maior percentual de ordens atrasadas e atraso médio, porque as ordens aguardaram mais tempo no *pre-shop pool* para serem liberadas. A dispersão dos resultados e os *outliers* também foram mais evidentes para os menores limites de WIP.

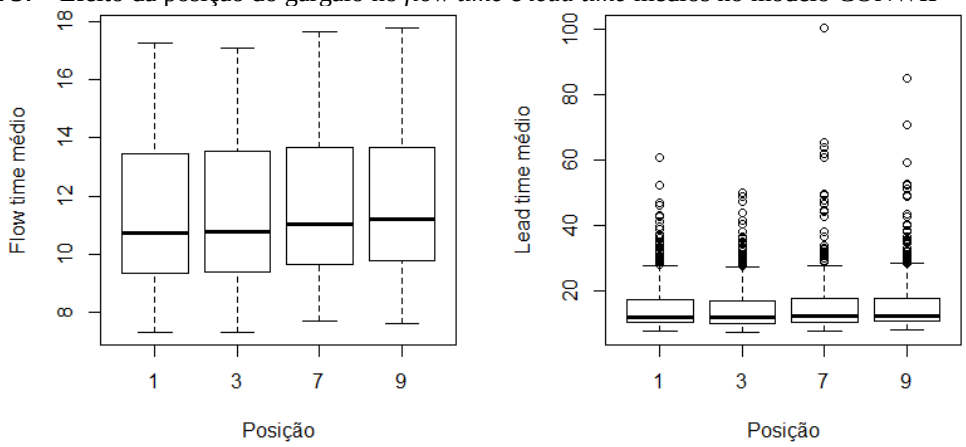
Figura 36 - Efeito dos níveis de maxWIP no percentual de ordens atrasadas e atraso médio no modelo CONWIP



Fonte: Elaboração própria.

No que diz respeito ao fator posição do gargalo nas simulações do CONWIP, verifica-se pouca diferença entre as posições no resultado do tempo médio de fluxo e *lead time* médio (Figura 37). Para as posições 1 e 3 a média do tempo médio de fluxo ficou em torno de 11,3 unidades de tempo e a mediana em 10,7 unidades, enquanto que para as posições 7 e 9 a média aumentou para 11,6 e 11,7, respectivamente e a mediana ficou um pouco acima das 11 unidades. O mesmo comportamento foi observado quanto ao *lead time*. As primeiras posições apresentaram média e mediana ligeiramente inferiores em relação as duas últimas localizações. Nos dois casos, houve maior dispersão dos resultados nas últimas máquinas.

Figura 37 - Efeito da posição do gargalo no *flow time* e *lead time* médios no modelo CONWIP

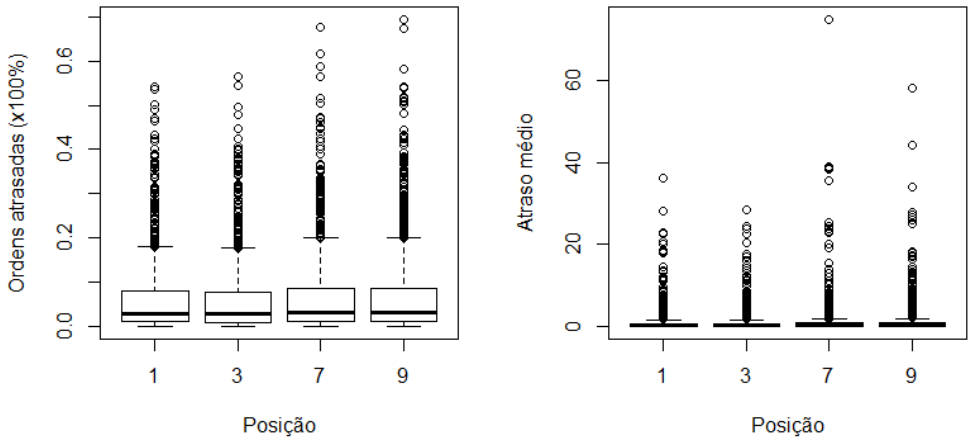


Fonte: Elaboração própria.

Em relação ao percentual de atraso, verifica-se que as duas primeiras posições apresentaram mediana próxima a 2% e média de 5%, enquanto que as localizações 7 e 9 obtiveram 3% e 6% para a média e a mediana, respectivamente. Essa diferença também foi identificada quanto ao indicador atraso médio. Quando o gargalo estava localizado na máquina 1 ou 3, o atraso médio foi de aproximadamente 0,8 unidades de tempo, já nas máquinas 7 e 9 o atraso apresentou média um pouco superior a 1 unidade de tempo (Figura 38).

Desse modo, nos quatro indicadores observou-se pequenas diferenças entre os resultados das máquinas 1 e 3 em relação ao gargalo nas máquinas 7 e 9, sendo a última localização com maior dispersão dos resultados e maiores valores médios.

Figura 38 - Efeito da posição do gargalo no percentual de ordens atrasadas e atraso médio no modelo CONWIP

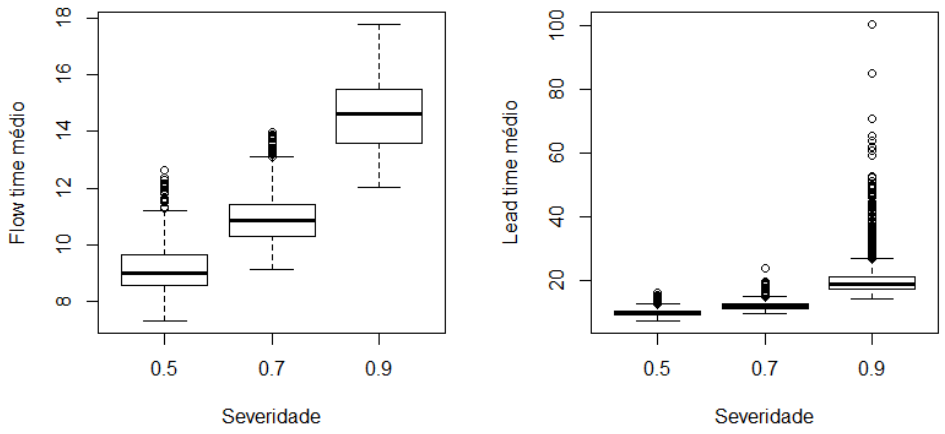


Fonte: Elaboração própria.

O efeito da severidade do gargalo pode ser observado nas Figuras 39 e 40. Quanto aos tempos médios de fluxo e de atravessamento dos *jobs*, verifica-se um aumento da média e mediana à medida que a severidade do gargalo diminui. Esse resultado era esperado, pois como comentado no Capítulo 5, para caracterizar o desbalanceamento, o tempo médio de processamento dos recursos não críticos foi reduzido em relação ao gargalo. Dessa forma, quanto maior a severidade do gargalo, menor a soma dos tempos de processamento e melhor o desempenho do sistema (THÜRER; STEVENSON, 2018).

As instancias com severidade baixa (0,9), onde a diferença entre o tempo médio de processamento no gargalo e nos recursos não-restritivos é menor, apresentaram maior dispersão dos resultados, especialmente, no *lead time* médio. O aumento da média dos resultados, tanto no *flow time* como no *lead time*, foi mais expressivo entre as severidades média (0,7) e baixa (0,9).

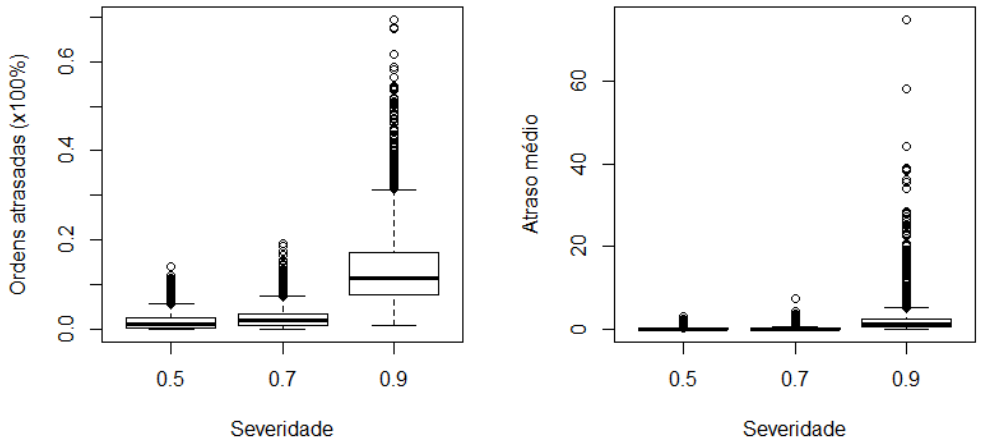
Figura 39 - Efeito da severidade do gargalo no *flow time* e *lead time* médios no modelo CONWIP



Fonte: Elaboração própria.

Quanto a confiabilidade de entrega, também se observa um comprometimento maior dos resultados quando a severidade do gargalo era baixa (0,9) como mostra a Figura 40. Em relação ao percentual de atraso, quando a severidade do gargalo era alta (0,5) ou média (0,7), o percentual médio oscilou entre 1,7% e 2,5%, respectivamente. Enquanto que nos cenários com severidade baixa, o percentual médio passou para 14,15%. Esse mesmo comportamento é observado no atraso médio, houve maior quantidade de *outliers* e um aumento do atraso das ordens quando a capacidade protetiva foi menor. Nas severidades média e alta, o atraso médio não chegou a uma unidade de tempo. Já nas simulações com severidade baixa a média passou para 2,5 unidades, uma vez que quando há menores níveis de capacidade protetiva nos recursos não críticos, há maiores chances de comprometimento do desempenho (BLACKSTONE JR; COX III, 2002).

Figura 40 - Efeito da severidade do gargalo no percentual de ordens atrasadas e atraso médio no modelo CONWIP



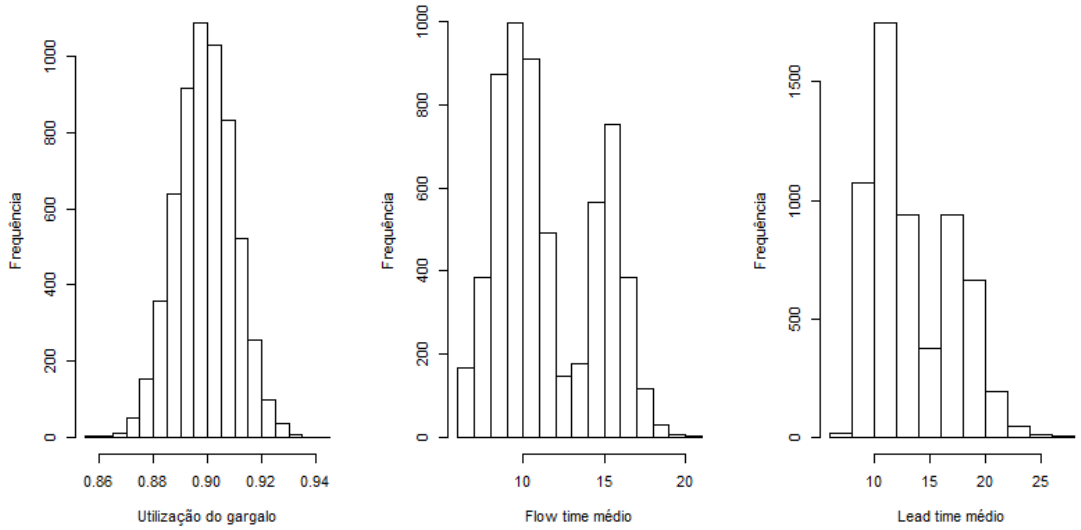
Fonte: Elaboração própria.

9.3 Resultados do DBR

De modo semelhante ao realizado na análise do CONWIP, inicialmente, foram elaborados histogramas para verificar a distribuição dos resultados dos cinco indicadores nas simulações do DBR (Figuras 41 e 42). Quanto a utilização do gargalo, observa-se também uma distribuição simétrica em torno de 90% de utilização.

Em relação ao tempo médio de fluxo, verifica-se que não há simetria nos resultados, havendo duas áreas com grande concentração dos casos (distribuição bimodal), próximo a 10 e a 15 unidades de tempo. A distribuição do *lead time* médio apresentou comportamento semelhante, havendo maior frequência dos resultados em torno de 12 unidades e 18 unidades de tempo (Figura 41).

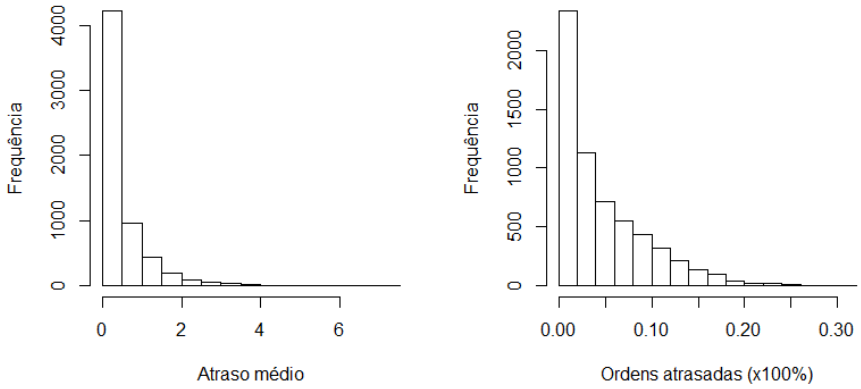
Figura 41 - Distribuição dos resultados dos indicadores utilização do gargalo, tempo médio de fluxo e *lead time* médio no DBR



Fonte: Elaboração própria.

Quanto aos indicadores de confiabilidade de entrega, observa-se que ambas as distribuições são assimétricas positivas, assim como verificado no CONWIP. O atraso médio apresentou maior concentração de resultados inferiores a 2 unidades de tempo e o percentual de ordens atrasadas na maioria das simulações foi inferior a 10% (Figura 42).

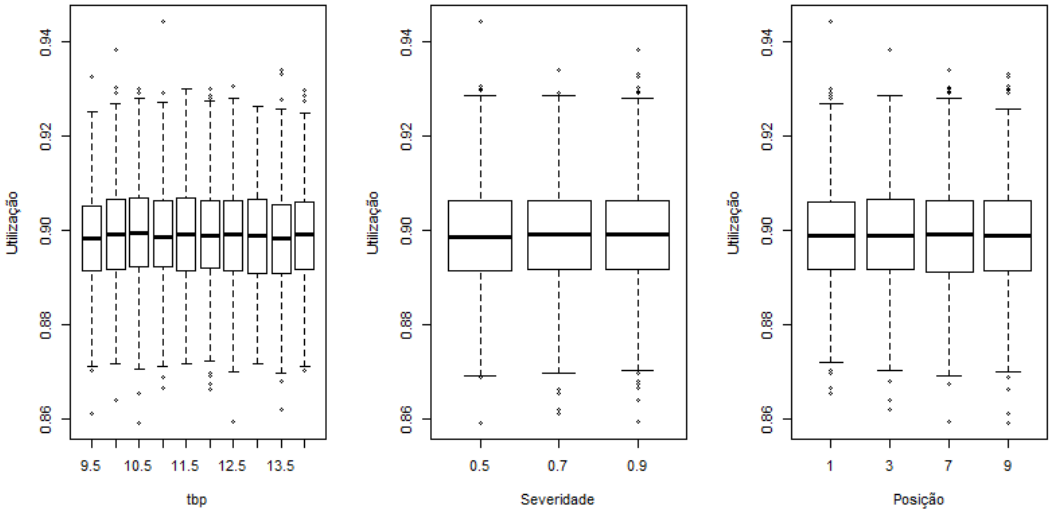
Figura 42 - Distribuição dos resultados dos indicadores atraso médio e percentual de ordens atrasadas no DBR



Fonte: Elaboração própria.

Ao estratificar os resultados da utilização média do gargalo pelos fatores do delineamento dos experimentos, observa-se na Figura 43 que de modo semelhante ao verificado no CONWIP, não há influência dos fatores na utilização. Para todos os níveis de *time buffer*, severidades e posições do gargalo, a taxa de utilização apresentou a média de 90%, restringida pela taxa de chegada dos *jobs* no sistema. Desse modo, os efeitos dos fatores no resultado das variáveis foram discutidos apenas para os demais indicadores do delineamento.

Figura 43 - Efeito dos fatores na utilização do gargalo no modelo DBR

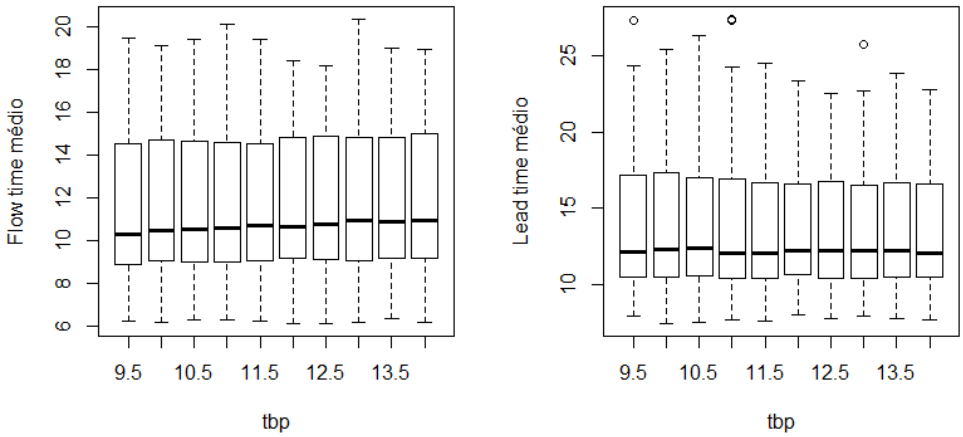


Fonte: Elaboração própria.

Inicialmente, pode-se analisar o efeito dos níveis de *time buffer* da produção (*tbp*) nos indicadores tempo médio de fluxo e *lead time* (Figura 44) e no percentual de atraso e atraso médio (Figura 45). Assim como ocorreu no CONWIP, observou-se que o aumento do *tbp*, aumentou a média do indicador *flow time* como sugere a literatura (KIM; COX; MABIN, 2010). No entanto, os incrementos foram menores e a dispersão dos resultados foi maior do que no

CONWIP. Em relação ao *lead time* médio, a análise dos *box plots* não evidenciou diferenças entre os níveis de *tbp*. A mediana e a média dos resultados foram semelhantes em todos os níveis.

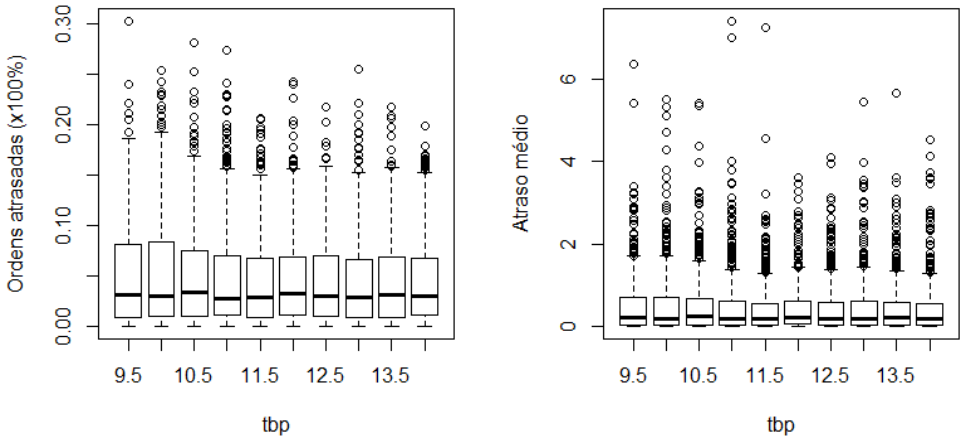
Figura 44 - Efeito do tbp no *flow time* e *lead time* médios no modelo DBR



Fonte: Elaboração própria.

A Figura 45 apresenta o efeito do *time buffer* na confiabilidade de entrega. Apesar de haver uma diminuição da dispersão dos resultados, não ficou evidente nos resultados dos experimentos que a adoção de mais proteção (maior *tbp*) contribuiu para uma melhoria no desempenho de entrega. O percentual médio oscilou em torno de 5% e o atraso médio por volta de 0,4 unidades de tempo.

Figura 45 - Efeito do tbp no percentual de ordens atrasadas e atraso médio no modelo DBR

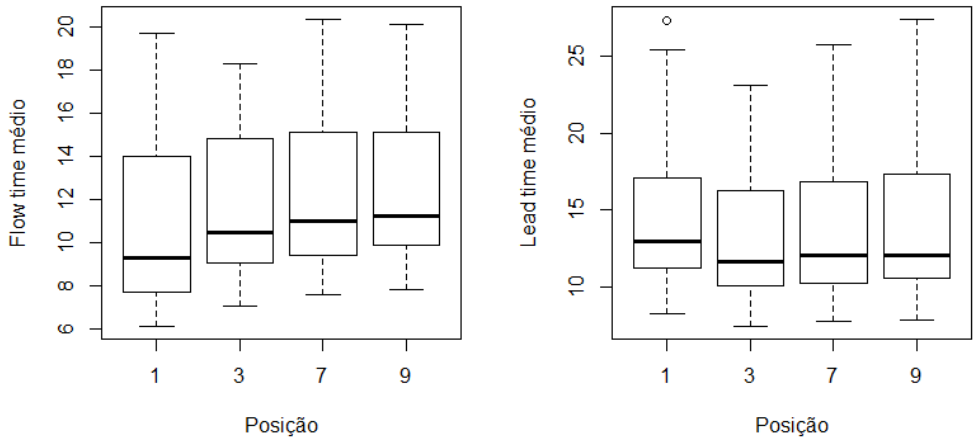


Fonte: Elaboração própria.

Em relação ao efeito da posição do gargalo no modelo DBR, a Figura 46 mostra os resultados para os indicadores tempo médio de fluxo e *lead time* médio. A média do tempo de fluxo aumentou quando a localização avaliada do gargalo era mais próxima do final do sistema.

A posição 1 apresentou maior dispersão dos resultados do tempo de fluxo, o que pode estar relacionado com a liberação condicionada apenas ao *backlog* do gargalo, já que não foi adicionada a proteção do gargalo em virtude do suprimento “infinito”. Quanto ao tempo médio de atravessamento, não foi possível observar um aumento sequencial da média ou mediana do indicador como no tempo de fluxo. A primeira posição apresentou maior valor médio, seguida pela localização 9, cenários que também apresentaram maior dispersão dos resultados, enquanto que a máquina 3 obteve melhor resultado.

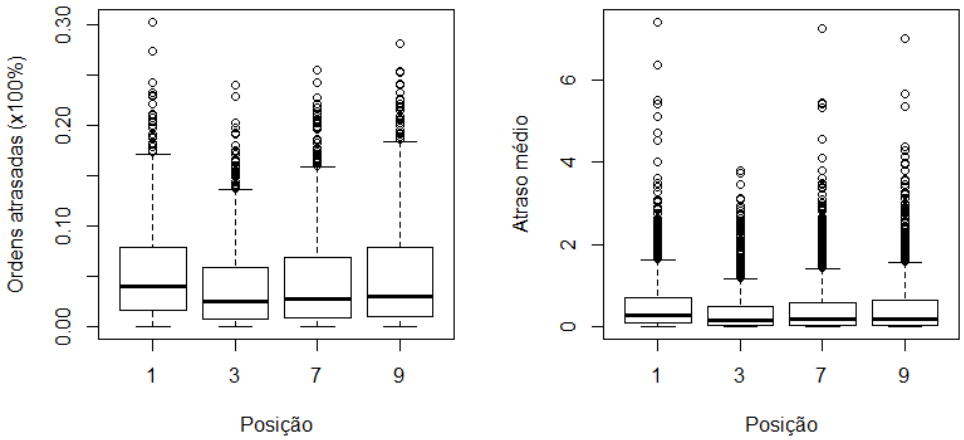
Figura 46 - Efeito da posição do gargalo no *flow time* e *lead time* médios no modelo DBR



Fonte: Elaboração própria.

Os resultados do percentual de ordens atrasadas e atraso médio apresentaram comportamento semelhante ao *lead time* médio. O maior valor médio do percentual e do atraso foram verificados quando o gargalo estava na posição 1, seguidos pelas instâncias com o gargalo na máquina 9. Os melhores resultados nesses indicadores foram quando o gargalo estava localizado na máquina 3 (Figura 47).

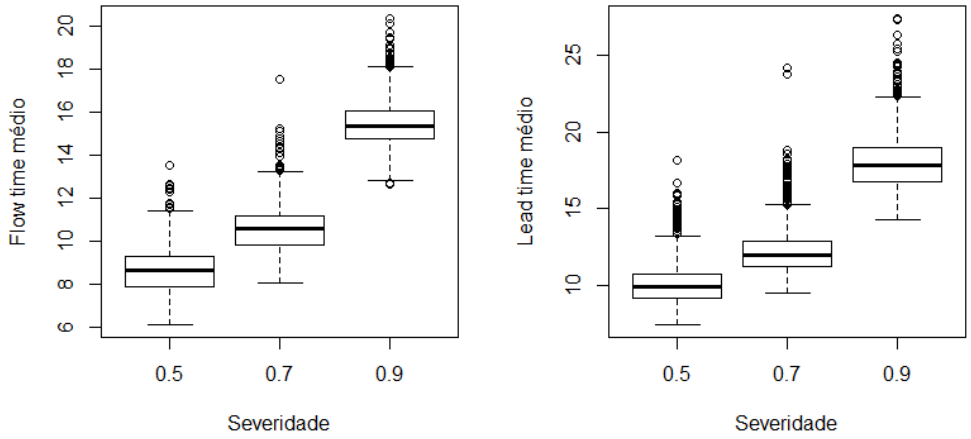
Figura 47 - Efeito da posição do gargalo no percentual de ordens atrasadas e atraso médio no modelo DBR



Fonte: Elaboração própria.

A severidade do gargalo apresentou efeito semelhante nas variáveis tempo médio de fluxo e *lead time* médio nos modelos CONWIP e DBR. Pode-se observar na Figura 48, que quando a severidade do gargalo foi menor, houve um aumento da média e mediana desses indicadores. Como mencionado anteriormente, esse resultado era esperado em virtude da configuração dos tempos médios de processamento das estações.

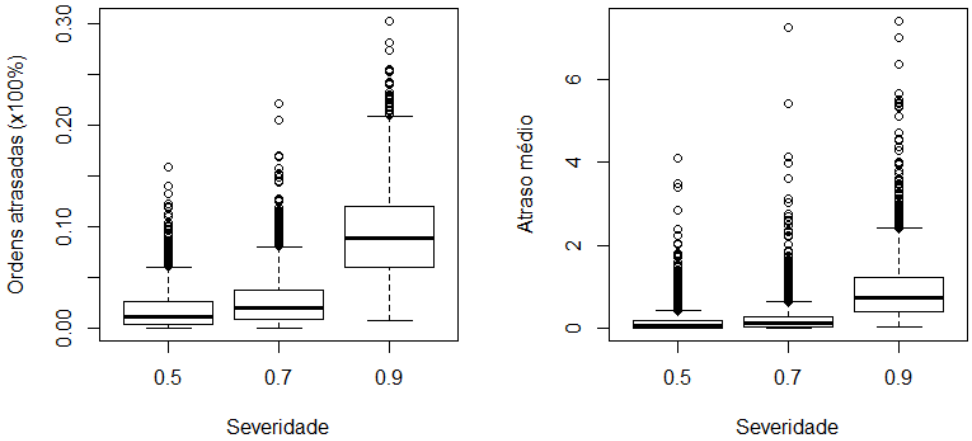
Figura 48 - Efeito da severidade do gargalo no *flow time* e *lead time* médios no modelo DBR



Fonte: Elaboração própria.

O efeito da severidade do gargalo também se evidencia nos indicadores: percentual de atraso e atraso médio (Figura 49). Nos cenários com severidade baixa (0,9) houve maior comprometimento do desempenho de entrega e maior dispersão dos resultados, assim como no CONWIP. O percentual médio de atraso das ordens foi de 1,8% na severidade alta, 2,7% na média e 9,4% na baixa. Em relação ao atraso médio, o valor médio das observações foi de 0,16; 0,24 e 0,9 unidades de tempo nos cenários com severidade alta, média e baixa, respectivamente.

Figura 49 - Efeito da severidade do gargalo no percentual de ordens atrasadas e atraso médio no modelo DBR



Fonte: Elaboração própria.

9.4 Análise de Variância

Com o objetivo de confirmar o efeito dos fatores dos experimentos e das interações desses no desempenho das variáveis de resposta, realizou-se uma Análise de Variância (ANOVA). O experimento deste estudo pode ser classificado como hierárquico-cruzado, uma vez que os níveis de *buffer* são diferentes para cada modelo de programação e controle (hierárquico) e são feitas todas as combinações entre os fatores (fatorial completo, experimento cruzado) (MONTGOMERY, 2017).

Por se tratar de um estudo teórico de simulação, há menor variabilidade dos dados em relação a um problema real. Nesse sentido, optou-se por considerar respostas médias dos resultados das 50 rodadas (tratamentos não replicados) a fim de aumentar o poder de detecção da análise. No entanto, quando não há réplicas no experimento, não é possível realizar a ANOVA, pois é preciso julgar se as estimativas observadas são razoáveis. Logo, como em um experimento fatorial, os efeitos principais e as interações de baixa ordem são significativos com maior frequência, enquanto que interações de alta ordem (3 ou mais fatores) raramente geram efeitos estatisticamente significativos, adotou-se a estratégia de utilizar essas interações de alta ordem para a estimativa do erro (pseudo-resíduo) (DANIEL, 2008; MONTGOMERY, 2017).

Além disso, Montgomery (2017) afirma que experimentos aleatorizados dispensam a suposição da normalidade dos dados. Por essa razão, não foram realizados testes de normalidade para a aplicação da ANOVA.

Como discutido anteriormente, o indicador utilização do recurso gargalo não apresentou variação com os fatores dos experimentos, estando condicionado ao parâmetro taxa de chegada. Dessa forma, a ANOVA foi realizada apenas para quatro indicadores: tempo médio de fluxo, *lead time* médio, atraso médio e percentual de atraso.

Para a variável de resposta tempo médio de fluxo, os efeitos principais dos fatores modelo, posição e severidade se mostraram estatisticamente significativos. Além disso, as interações entre os fatores "*Buffer* x Severidade" e "*Buffer* x Posição" se mostraram significativos a um nível de 0,1%, enquanto que as interações "Modelo x Posição" e "Modelo x Severidade", foram estatisticamente significativos a 5% de significância. A Tabela 8 apresenta os resultados para a variável tempo médio de fluxo.

Tabela 8 - ANOVA da variável de resposta tempo médio de fluxo

Tempo médio de fluxo	Graus de liberdade	Soma dos quadrados	Quadrado médio	F	Pr(>F)
Modelo	1	26112747	26112747	19,2962	2,374e-05 ***
Posição	3	29750679	9916893	7,3282	0,0001457 ***
Severidade	2	887824043	443912022	328,0321	< 2,2e-16 ***
Modelo:*Buffer*	18	26365893	1464772	1,0824	0,3775054
Modelo:Posição	3	17478050	5826017	4,3052	0,0063188 **
Modelo:Severidade	2	19163094	9581547	7,0804	0,0012255 **
***Buffer*:Posição**	48	10259842	213747	0,1579	1,0000000
***Buffer*:Severidade**	32	145731733	4554117	3,3653	7,228e-07 ***
Posição:Severidade	6	136283593	22713932	16,7846	4,132e-14 ***
Erro	124	167803974	1353258		

Legenda: *** probabilidade significativa entre 0 e 0,1%; ** probabilidade significativa entre 0,1 e 1%.
Fonte: Elaboração própria.

Quanto ao indicador *lead time* médio, observou-se que ao considerar um nível de significância de 5%, há evidências da influência dos fatores modelo e severidade do gargalo e das interações entre os fatores “Modelo x Severidade” e “Posição x Severidade”. O fator posição do gargalo mostrou efeito significativo do ponto de vista estatístico somente ao considerar 10% de significância, como apresentado na Tabela 9.

Tabela 9 - ANOVA da variável de resposta *lead time* médio

***Lead time* médio**	Graus de Liberdade	Soma dos quadrados	Quadrado médio	F	Pr(>F)
Modelo	1	4413368	4413368	9,2170	0,002923 **
Posição	3	3570025	1190008	2,4852	0,063797 ‘.’
Severidade	2	792364763	396182382	827,3967	< 2,2e-16 ***
Modelo:*Buffer*	18	9650729	536152	1,1197	0,341346
Modelo:Posição	3	2846940	948980	1,9819	0,120143
Modelo:Severidade	2	5417790	2708895	5,6573	0,004454 **
***Buffer*:Posição**	48	7714011	160709	0,3356	0,999979
***Buffer*:Severidade**	32	13600941	425029	0,8876	0,641748
Posição:Severidade	6	28146472	4691079	9,7970	7,952e-09 ***
Erro	124	59374921	478830		

Legenda: *** probabilidade significativa entre 0 e 0,1%; ** probabilidade significativa entre 0,1 e 1%; ‘.’ probabilidade significativa a 10%.
Fonte: Elaboração própria.

As variáveis de resposta atraso médio e percentual de atraso das ordens apresentaram comportamento semelhante. Nesses casos, todos os fatores e as interações, exceto a interação dos fatores “*Buffer* x Posição”, influenciaram os resultados dessas variáveis. As Tabelas 10 e 11 apresentam os resultados para essas variáveis de resposta.

Tabela 10 - ANOVA da variável de resposta atraso médio

Atraso médio	Graus de liberdade	Soma dos quadrados	Quadrado médio	F	Pr(>F)
Modelo	1	4329633	4329633	27,3828	6,896e-07 ***
Posição	3	10459062	3486354	22,0495	1,647e-11 ***
Severidade	2	1306486491	653243245	4131,4448	1,490e-06 ***
Modelo:*Buffer*	18	11526781	640377	4,0501	1,490e-06 ***
Modelo:Posição	3	7232203	2410734	15,2467	1,664e-08 ***
Modelo:Severidade	2	10585786	5292893	33,4750	2,372e-12 ***
***Buffer*:Posição**	48	7195095	149898	0,9480	0,5733559
***Buffer*:Severidade**	32	26792479	837265	5,2953	7,192e-12 ***
Posição:Severidade	6	3947892	657982	4,1614	0,0007612 ***
Erro	124	19606255	158115		

Legenda: *** probabilidade significativa entre 0 e 0,1%.
Fonte: Elaboração própria.

Tabela 11 - ANOVA da variável de resposta percentual de atraso

Percentual de atraso	Graus de liberdade	Soma dos quadrados	Quadrado médio	F	Pr(>F)
Modelo	1	2851401	2851401	21,5965	8,464e-06 ***
Posição	3	14026028	4675343	35,4110	< 2.2e-16 ***
Severidade	2	1584016365	792008183	5998,6682	< 2.2e-16 ***
Modelo:*Buffer*	18	9459749	525542	3,9805	2,037e-06 ***
Modelo:Posição	3	10784079	3594693	27,2262	1,350e-13 ***
Modelo:Severidade	2	9373347	4686673	35,4969	6,469e-13 ***
***Buffer*:Posição**	48	4999145	104149	0,7888	0,8239016
***Buffer*:Severidade**	32	20974212	655444	4,9643	4,694e-11 ***
Posição:Severidade	6	3972919	662153	5,0151	0,0001237 ***
Erro	124	16371803	132031		

Legenda: *** probabilidade significativa entre 0 e 0,1%.
Fonte: Elaboração própria.

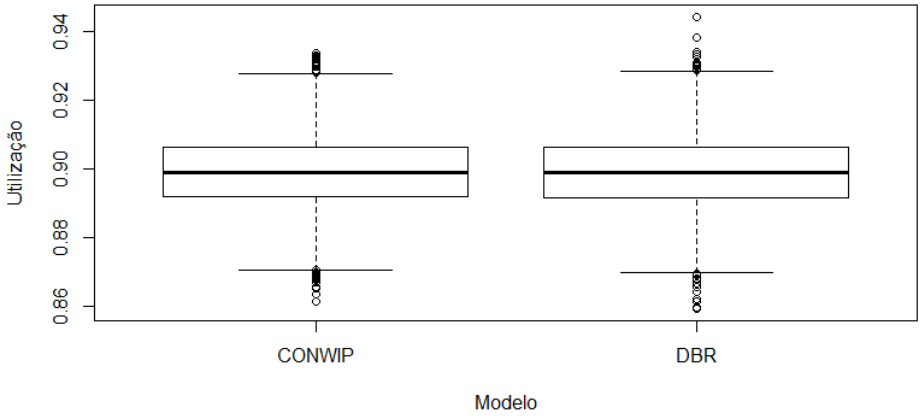
Desse modo, ao considerar 5% de significância verifica-se que os fatores: modelo de programação e controle e severidade do gargalo afetam todos os indicadores. O fator posição do gargalo também influencia todas as variáveis de resposta, mas para o *lead time* médio deve-se considerar 10% de significância. Para todos os indicadores, a interação entre os fatores "*Buffer* x Posição" não se mostrou significativa, ou seja, a mudança dos dois fatores conjuntamente não influencia o desempenho. Uma das possíveis razões para essa evidência, é o fato de nos experimentos do DBR o nível de proteção (*tbp*) ser ajustado de acordo com a posição do gargalo, multiplicado pelas constantes para cálculo do *time buffer* do gargalo.

9.5 Comparação dos resultados dos modelos: CONWIP x DBR

Os resultados dos modelos CONWIP e DBR foram comparados frente ao desempenho nos indicadores. A Figura 50 apresenta os resultados dos modelos frente a utilização do recurso gargalo. Como observado nas seções anteriores, os dois sistemas tiveram desempenho

semelhante, garantindo boa utilização do gargalo. Desse modo, não há evidências de que o modelo adotado foi determinante para a utilização média do recurso crítico.

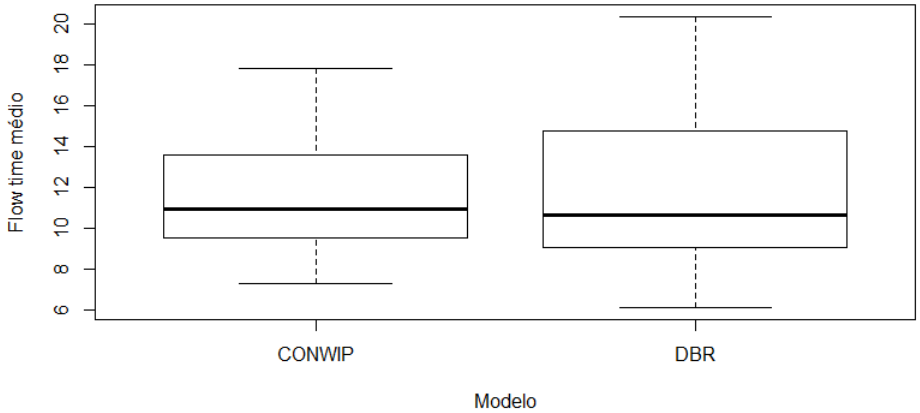
Figura 50 - Desempenho dos modelos CONWIP e DBR quanto a utilização do gargalo



Fonte: Elaboração própria.

Na avaliação do tempo médio de fluxo, observa-se algumas diferenças entre os modelos de programação e controle, como mostra a Figura 51. O CONWIP e o DBR apresentaram valor médio semelhante (11,2 unidades de tempo) e medianas muito próximas. O DBR, no entanto, teve uma variação maior dos resultados em relação ao CONWIP com valor mínimo e máximo de 6,12 e 20,33 unidades, respectivamente. Enquanto que no CONWIP, o valor mínimo foi 7,3 e o valor máximo foi 17,78 unidades de tempo.

Figura 51 - Desempenho dos modelos CONWIP e DBR quanto ao tempo médio de fluxo

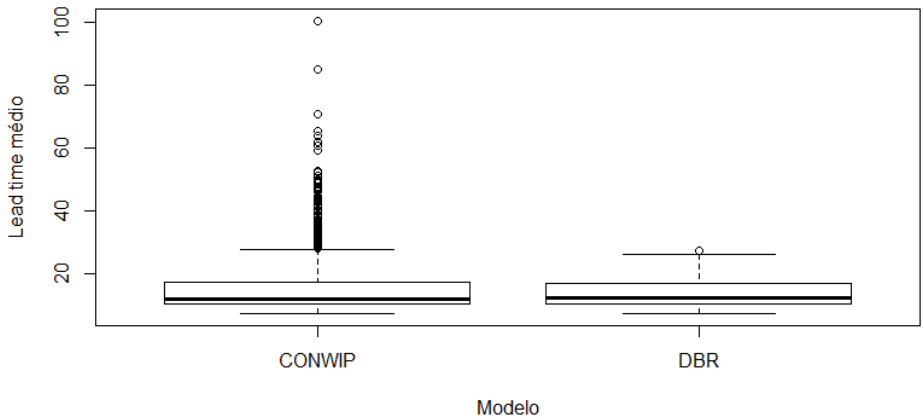


Fonte: Elaboração própria.

Em relação ao *lead time* médio, os modelos apresentaram valor médio próximo a 14 unidades de tempo. O valor mínimo também foi semelhante nos dois modelos, contudo, o

CONWIP apresentou mais valores discrepantes, com valor máximo próximo a 100 unidades de tempo (Figura 52).

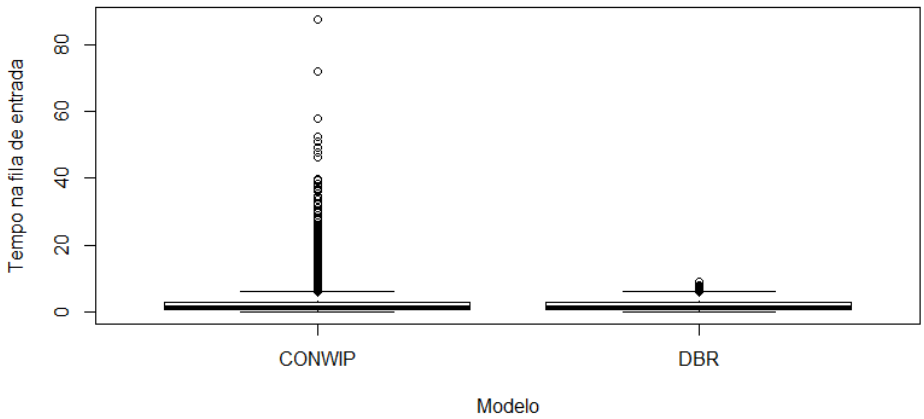
Figura 52 - Desempenho dos modelos CONWIP e DBR quanto ao *lead time* médio



Fonte: Elaboração própria.

A diferença entre os indicadores *lead time* e tempo de fluxo é dada pelo tempo que a ordem aguarda no *pre-shop pool*. A Figura 53 apresenta os tempos médios que as ordens esperaram para liberação. Verifica-se que no modelo CONWIP, houve maior variação dos tempos, de modo que algumas ordens esperaram mais do que 80 unidades de tempo para liberação. Os maiores tempos de espera registrados foram de instâncias com severidade baixa do gargalo (menor desbalanceamento das operações e maior tempo total de processamento) e menores níveis de maxWIP (controle mais restritivo). No DBR, o valor máximo do tempo na fila foi de aproximadamente 9 unidades de tempo.

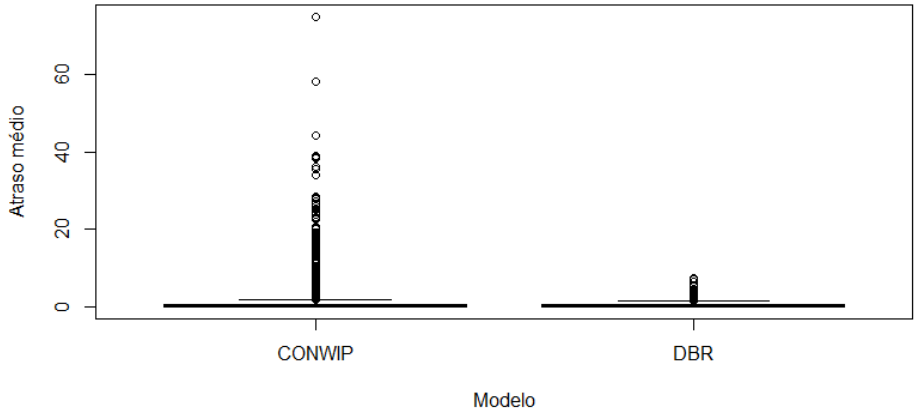
Figura 53 - Tempo médio de espera na fila de entrada com os modelos CONWIP e DBR



Fonte: Elaboração própria.

Quanto ao indicador atraso médio, ambos os modelos apresentaram resultados baixos. No caso do CONWIP, houve uma grande concentração de casos com atrasos inferiores a 1 unidade de tempo (média de 0,95 unidades), mas com bastante *outliers* (valores superiores até a 60 unidades de tempo). O DBR apresentou resultados de atrasos mais concentrados com média menor que a do CONWIP (0,44 unidades de tempo) (Figura 54).

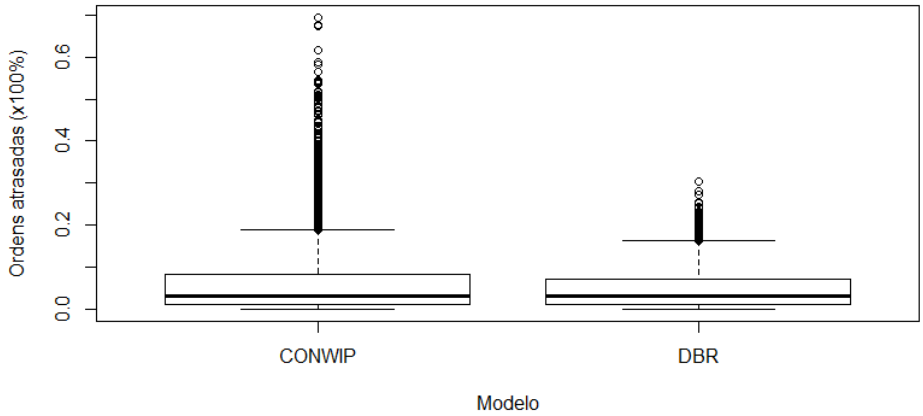
Figura 54 - Desempenho dos modelos CONWIP e DBR quanto ao atraso médio



Fonte: Elaboração própria.

Ao analisar os resultados do percentual de ordens atrasadas estratificados pelos modelos na Figura 55, verifica-se que a mediana da variável foi semelhante nos dois modelos. Já média do percentual foi de 6,1% no CONWIP e 4,6% no DBR. O CONWIP, novamente, apresentou maior dispersão dos resultados com máximo superior a 60% das ordens atrasadas, enquanto que no DBR o percentual máximo foi de 30% de atraso.

Figura 55 - Desempenho dos modelos CONWIP e DBR quanto ao percentual de ordens atrasadas



Fonte: Elaboração própria.

Para verificar se existe diferença estatística entre os modelos, realizou-se um teste *t* de *Student* com as respostas dos indicadores de desempenho. Para aplicação do teste *t*, assume-se que as duas amostras em comparação são aleatórias, independentes e podem ser descritas com uma distribuição normal. Montgomery (2017) discute que pequenos desvios da distribuição normal não afetam os resultados e defende que experimentos aleatorizados permitem o teste de hipóteses sem consideração quanto a forma da distribuição. Como o experimento simulado foi aleatorizado e a distribuição baseada na aleatorização se aproxima da distribuição *t* de *Student*, esse teste foi aplicado para evidenciar a diferenças entre as médias das variáveis de resposta.

As hipóteses nulas testadas (H_0) são de que a média de cada uma das respostas das variáveis do modelo CONWIP (tempo médio de fluxo, *lead time* médio, percentual de atraso e atraso médio) são iguais as médias das respostas no DBR. Caso contrário, tem-se as hipóteses alternativas (H_1): as médias das respostas do CONWIP são diferentes das médias das variáveis de resposta do DBR.

Os resultados são apresentados na Tabela 12, considerando um nível de significância α=0,05. O valor *t* representa a estatística do teste calculado, comparado com a distribuição *t* de *Student* com 238 graus de liberdade (tamanho da amostra do CONWIP - 1 + tamanho da amostra do DBR - 1). O valor P é a probabilidade de ser obter um valor tão extremo como *t* na distribuição *t*. A hipótese nula é rejeitada no caso em que o valor P for igual ou inferior a 0,05 e aceita caso contrário (TRIOLA, 2005).

Como pode-se observar na Tabela 12, há indícios de diferença entre os modelos quanto ao indicador tempo médio de fluxo. O CONWIP apresentou média inferior ao DBR, o que sugere uma efetividade maior do modelo CONWIP quanto ao controle desse indicador. Nas demais variáveis de resposta, não se observou diferença entre os modelos, sugerindo similaridade no desempenho desses sistemas de programação e controle.

Tabela 12 - Resultado da comparação dos modelos com teste t

Variáveis de resposta	***t***	**Valor P**	**Conclusão**
Tempo médio de fluxo	2,077	0,03888	Rejeita H_0
***Lead time* médio**	1,067	0,2871	Aceita H_0
Atraso médio	0,85675	0,3924	Aceita H_0
Percentual de atraso	0,63671	0,5249	Aceita H_0

Fonte: Elaboração própria.

9.6 Discussões

As discussões dos resultados da pesquisa foram segmentadas de acordo com as questões que norteiam este trabalho: (1) Qual é o impacto da posição do gargalo no desempenho do controle efetuado pelo CONWIP e DBR em um ambiente de montagem MTO?; (2) Qual modelo é mais apropriado para o controle de um *assembly flow shop* que produz contra pedido?

9.6.1 Questão de pesquisa 1

A primeira questão de pesquisa deste estudo remete ao impacto da posição do gargalo no desempenho dos modelos de programação e controle CONWIP e DBR no ambiente de montagem MTO. Como observado na ANOVA, o fator posição influencia o desempenho dos indicadores tempo médio de fluxo, *lead time* médio, percentual de atraso e atraso médio.

Como discutido na revisão da literatura, o DBR clássico depende da localização do gargalo para programação e controle da entrada das ordens. Enquanto que o CONWIP, em sua versão original, não depende dessa informação para executar o controle da liberação e o sequenciamento ou programação da fila de entrada. Desse modo, houve diferenças no desempenho dos modelos diante da posição do gargalo.

Observou-se pouca variação nos resultados dos indicadores entre as diversas posições do gargalo no modelo CONWIP. De modo geral, as posições 1 e 3 apresentaram resultado ligeiramente superior as duas últimas (máquinas 7 e 9). Essas constatações corroboram com Thürer et al. (2017b), os autores afirmam que quanto mais ao início do processo estiver localizado o gargalo, mais efetivo o controle do sistema aplicado.

Esse resultado complementa as considerações de Duenyas (1994a), que havia observado maior comprometimento da vazão quando o gargalo era a montagem em relação ao gargalo nas linhas de fabricação em um sistema com montagem operado com CONWIP. Como também, confirma parcialmente as observações de Fry et al. (1987), de que quando o gargalo está localizado no início do processo de fabricação o sistema de montagem apresenta melhor desempenho.

Contudo, diferentemente do observado por Fry et al. (1987) no *assembly shop*, o pior desempenho do sistema não foi em operações anteriores a montagens, mas quando esse estava localizado na estação de montagem ou na última operação do sistema. Como esse estudo não considerou o controle da produção, houve maior dificuldade na coordenação das operações e sincronização da montagem.

As constatações da presente pesquisa também divergem de Kim; Davis e Cox III (2003), que verificou melhor desempenho do CONWIP quando a restrição estava localizada no final da linha. Possivelmente, essas diferenças são decorrentes da configuração dos experimentos, que considerou apenas um nível de severidade do gargalo e quebras das máquinas em diferentes posições do *flow shop*.

A fim de ampliar as investigações do impacto da posição do gargalo nos modelos e assumindo que a análise da posição pode ajudar a explicar o efeito da flutuação no desempenho de ambientes com controle da produção (THÜRER et al., 2017b; THÜRER; STEVENSON, 2018), elaborou-se a Tabela 13.

Na Tabela 13 são estratificados os resultados dos modelos de acordo com a posição do gargalo (g) e o nível de *buffer* (diferentes para cada modelo). Os valores apresentados são as médias de 150 instâncias, que contemplam os três níveis de severidade e 50 replicações de cada problema. Os resultados referentes as instâncias com o gargalo na máquina 1 estão sombreados em cinza. As colunas seguintes são destacadas em verde quando verificou-se uma redução do valor do indicador em relação a posição anterior e em vermelho quando o resultado da variável aumentou (o desempenho piorou).

Ao analisar os resultados médios do CONWIP na Tabela 13, verifica-se que para as variáveis atraso médio e percentual de atraso, a primeira e a terceira posição apresentaram desempenhos semelhantes. Quando o gargalo estava na montagem (g=7), houve um comprometimento mais expressivo do desempenho médio. Os resultados da última máquina do sistema (g=9), em algumas situações, foram melhores e em outras piores do que quando localizado na montagem. Para o percentual de atraso, as diferenças foram mínimas.

Em relação ao tempo médio de fluxo, evidencia-se nos resultados um comprometimento maior do desempenho do CONWIP quando o gargalo estava mais próximo ao final do sistema (g=7 e g=9). Os resultados do *lead time* médio apresentaram comportamento semelhantes ao *flow time*, contudo, houve dois casos em que o desempenho melhorou quando o gargalo passou a ser a máquina 9.

Como verificado inicialmente na seção 9.2, não foi possível verificar diferenças expressivas entre o desempenho das instâncias com o gargalo localizado nas máquinas 1 e 3. Para todos os indicadores, na maioria dos casos, os resultados médios das instâncias com o gargalo na última máquina foram maiores do que quando o gargalo estava na primeira máquina.

No CONWIP, como o controle do WIP é para toda a linha, esse modelo não é drasticamente afetado com mudança da posição do gargalo. Como pode-se observar na Tabela 12, a medida que o maxWIP aumenta, as diferenças entre as médias dos indicadores nas

diversas posições do gargalo diminuem. Assim, pode-se complementar as observações de Thürer et al. (2017b), quando os níveis de WIP são menores, o gargalo localizado nas primeiras estações do sistema contribui para um melhor desempenho e controle mais efetivo do sistema.

Em relação ao impacto da posição do gargalo no DBR, observou-se que os indicadores *lead time* médio, atraso médio e percentual de atraso apresentaram comportamento semelhantes: piores desempenhos nas máquinas 1 e 9 (extremos do sistema), melhor desempenho na posição 3 (última máquina da linha de fabricação), seguida pela montagem com desempenho superior a primeira e a última posição. Quanto ao tempo médio de fluxo, os resultados evidenciaram que a média desse indicador aumentou quando a localização avaliada do gargalo era mais próxima do final do sistema.

Thürer et al. (2017b) também verificaram que DBR teve uma perda do desempenho quando o gargalo estava ao final do sistema (*flow shop*). Além disso, esses resultados sugerem que, como discutido por Thürer e Stevenson (2018), o efeito da mudança do gargalo para uma estação posterior no sistema compromete mais o desempenho do DBR em termos de tempo de fluxo, seja o ambiente um *flow shop* ou um *assembly flow shop*.

O efeito da posição do gargalo no DBR observado nesta pesquisa foi diferente mais uma vez de Kim; Davis e Cox III (2003), que verificou melhor desempenho do DBR quando o gargalo estava no meio ou no final da linha e de Fry et al. (1987), evidenciando as distinções entre um sistema de montagem com e sem controle da liberação de ordens. O resultado desta pesquisa também divergiu de Lee e Seo (2016), que observaram melhor desempenho quando o gargalo estava posicionado no final da linha. No entanto, esse estudo não considerou aleatoriedade dos tempos de processamento.

Pode-se ratificar a afirmativa de trabalhos anteriores de que o DBR se assemelha ao CONWIP quando o gargalo está na última posição (SPEARMAN; WOODRUFF; HOPP, 1990; KIM; DAVIS; COX III, 2003; THÜRER; STEVENSON, 2018). Como observa-se na Tabela 13, apesar das variações dos valores, nos dois modelos houve um comprometimento do desempenho das variáveis de resposta quando o gargalo estava na última máquina.

É importante frisar que este trabalho, diferentemente dos estudos comparativos anteriores apresentados no Quadro 6, se propôs a simular o DBR aplicando o conceito de *time buffers* e não de estoque físico (controle da quantidade de ordens em processamento até o gargalo). Buscou-se representar com maior proximidade a lógica do DBR, como descrito na literatura clássica do modelo, e mitigar semelhanças do DBR com as variações do CONWIP com *loops* até o gargalo (PRAKASH; CHIN, 2015).

Tabela 13 - Impacto da posição do gargalo nos modelos CONWIP e DBR

Modelo	*Buffer*	Atraso médio				Percentual de atraso				Tempo médio de fluxo				*Lead time* médio			
		g=1	g=3	g=7	g=9	g=1	g=3	g=7	g=9	g=1	g=3	g=7	g=9	g=1	g=3	g=7	g=9
CONWIP	**12**	2,753	2,809	4,027	4,276	0,112	0,113	0,124	0,137	10,110	10,134	10,321	10,394	16,786	16,916	18,529	19,076
CONWIP	**13**	1,458	1,189	1,502	1,874	0,079	0,072	0,087	0,094	10,500	10,507	10,701	10,780	14,955	14,573	15,450	15,941
CONWIP	**14**	0,763	0,869	1,002	0,807	0,059	0,061	0,071	0,065	10,832	10,840	11,078	11,104	13,916	14,083	14,656	14,378
CONWIP	**15**	0,637	0,605	0,847	0,873	0,056	0,053	0,064	0,064	11,146	11,080	11,327	11,433	13,713	13,545	14,238	14,401
CONWIP	**16**	0,558	0,546	0,697	0,537	0,048	0,049	0,057	0,054	11,375	11,339	11,616	11,741	13,458	13,420	13,980	13,948
CONWIP	**17**	0,551	0,502	0,504	0,602	0,051	0,044	0,052	0,057	11,622	11,529	11,892	12,117	13,469	13,230	13,740	14,179
CONWIP	**18**	0,503	0,532	0,527	0,490	0,044	0,049	0,050	0,050	11,750	11,813	12,137	12,192	13,241	13,404	13,768	13,791
CONWIP	**19**	0,408	0,426	0,518	0,576	0,044	0,045	0,048	0,055	12,019	12,019	12,135	12,551	13,224	13,254	13,498	14,083
CONWIP	**20**	0,455	0,356	0,460	0,435	0,047	0,040	0,047	0,045	12,194	12,077	12,423	12,500	13,317	13,034	13,550	13,591
CONWIP	**21**	0,369	0,373	0,407	0,451	0,042	0,042	0,043	0,047	12,227	12,272	12,490	12,569	13,060	13,114	13,371	13,533
DBR	**9,5**	0,570	0,390	0,475	0,537	0,057	0,041	0,049	0,054	10,553	11,240	11,605	11,736	14,185	12,982	13,405	13,749
DBR	**10**	0,558	0,381	0,556	0,562	0,055	0,041	0,052	0,056	10,529	11,293	11,799	11,992	14,112	12,920	13,588	13,936
DBR	**10,5**	0,540	0,367	0,451	0,593	0,056	0,039	0,047	0,055	10,521	11,409	11,837	12,072	14,083	12,957	13,440	13,875
DBR	**11**	0,491	0,365	0,478	0,511	0,051	0,038	0,045	0,050	10,430	11,434	11,912	11,997	13,926	12,865	13,408	13,557
DBR	**11,5**	0,458	0,315	0,442	0,401	0,050	0,037	0,044	0,044	10,409	11,486	11,954	12,031	13,900	12,836	13,335	13,381
DBR	**12**	0,550	0,369	0,354	0,423	0,058	0,039	0,040	0,046	10,595	11,593	11,947	12,123	14,241	12,918	13,164	13,401
DBR	**12,5**	0,513	0,282	0,394	0,474	0,053	0,034	0,044	0,049	10,485	11,526	12,018	12,363	14,058	12,692	13,171	13,621
DBR	**13**	0,484	0,287	0,444	0,478	0,051	0,035	0,044	0,048	10,428	11,629	12,134	12,394	13,948	12,747	13,246	13,552
DBR	**13,5**	0,475	0,337	0,353	0,503	0,051	0,036	0,039	0,051	10,470	11,642	12,061	12,553	14,031	12,729	13,014	13,673
DBR	**14**	0,431	0,429	0,379	0,427	0,048	0,040	0,042	0,046	10,309	11,878	12,284	12,505	13,738	13,002	13,240	13,468

Fonte: Elaboração própria.

9.6.2 Questão de pesquisa 2

A segunda questão de pesquisa remete às diferenças entre o desempenho dos modelos no ambiente *assembly flow shop* MTO. Como observado na seção 9.5, o teste de hipóteses *t* de *Student* evidenciou diferença estatística entre os modelos quanto ao tempo médio de fluxo.

O resultado sugere que o CONWIP tem um controle mais efetivo do estoque em processo, corroborando com Jaegler et al. (2017), quanto ao desempenho superior do CONWIP frente a outros modelos em relação ao *flow time*. Assim como em Thürer et al. (2017b), na análise de um *flow shop*, os modelos DBR e CONWIP se mostraram estatisticamente semelhantes quanto ao desempenho dos indicadores *lead time* médio, atraso médio e percentual de ordens atrasadas.

Desse modo, verifica-se uma vantagem do CONWIP em relação ao DBR. Já que os modelos apresentaram desempenho semelhante e a literatura destaca a dependência do DBR, em a cada mudança do gargalo, montar uma nova programação do recurso e ajustar os controles (GILLAND, 2002; THÜRER et al., 2017b).

Esse resultado diverge de outras pesquisas que identificaram o DBR como superior ao CONWIP (GILLAND, 2002; KOH; BULFIN, 2004; NG; BERNEDIXEN; SYBERFELDT, 2012). Em nenhum desses casos, no entanto, o DBR foi simulado considerando *time buffers* (aplicaram o controle de WIP até o gargalo), a programação finita do gargalo e o gerenciamento de pulmões. Assim como, o CONWIP não foi configurado com o controle dinâmico de cartões.

Framinan et al. (2003) destacou a importância dos estudos comparativos com o CONWIP observarem o impacto da posição do gargalo (discutido anteriormente) e da diferença entre a capacidade do gargalo e a dos demais recursos (capacidade protetiva). Essa diferença pode ser discutida em função dos níveis de severidade adotados nesta pesquisa.

Diferentemente de Thürer et al. (2017b), nesse estudo, os níveis de severidade influenciaram o desempenho dos sistemas e evidenciaram diferenças entre o CONWIP e o DBR. Ao investigar os resultados do estudo, verificou-se que quando a severidade do gargalo foi baixa, o CONWIP apresentou menor valor médio para os indicadores tempo de médio de fluxo e percentual de atraso. Enquanto que, nos níveis de severidade média e alta, o DBR teve melhor desempenho.

Esses resultados corroboram com a literatura, ao evidenciar o bom desempenho do DBR em cenários com severidade alta (THÜRER et al., 2017c). Como também, sugerem a aplicação do CONWIP em situações onde há maior carga no sistema e a coordenação das operações se torna mais crítica (BERTRAND; FRANSOO, 2002).

Para os indicadores *lead time* médio e atraso médio, o DBR apresentou menor valor médio em relação ao CONWIP nas instâncias de severidade baixa. Esse comprometimento do desempenho do CONWIP foi influenciado pelo controle mais acentuado da fila de entrada nos menores níveis de maxWIP. Nos demais casos, severidade média e alta, o CONWIP mostrou um resultado um pouco melhor.

CAPÍTULO 10 – CONSIDERAÇÕES FINAIS

Este capítulo apresenta a síntese das conclusões e principais contribuições deste estudo. Em seguida, são descritas as limitações e possíveis desdobramentos da pesquisa.

10.1 Conclusões e contribuições da pesquisa

A presente tese comparou o desempenho dos modelos de programação e controle CONWIP e DBR em um *assembly flow shop* com produção MTO. A comparação contemplou a avaliação do impacto da posição do gargalo, severidade da restrição e dos limites de *buffer* no desempenho dos modelos, diante dos indicadores atraso médio, percentual de atraso, tempo médio de fluxo, *lead time* médio e utilização do gargalo.

Os modelos CONWIP e DBR são amplamente discutidos na literatura de Planejamento e Controle da Produção. Em especial, observam-se estudos recentes sobre a aplicabilidade desses métodos em ambientes com produção contra pedido, onde o foco principal é o controle das operações para lidar com a alta variedade e baixo volume dos lotes.

Os dois modelos já haviam sido comparados em alguns estudos na literatura, sobretudo, com escopo teórico e por meio de simulações. Contudo, observou-se que estes trabalhos consideraram na maioria dos casos ambientes simples como *flow shops*. Além disso, boa parte dessas comparações não avaliou os modelos diante das características da produção MTO. Desse modo, a presente tese compara os modelos em um ambiente ainda pouco discutido e contribui com discussão sobre como o CONWIP e o DBR podem ser aplicados em sistemas de montagem contra pedido.

A comparação dos modelos de programação e controle ratifica e estende conclusões de outros estudos da literatura como Thürer et al. (2017b), ao observar que os modelos apresentaram desempenho estatisticamente semelhante quanto ao *lead time* médio, atraso médio e percentual de atraso no ambiente *assembly flow shop* em estudo. Ou seja, para as métricas referentes ao desempenho de entrega (críticas para a produção MTO), os modelos se mostraram equivalentes. Em relação ao tempo médio de fluxo, no entanto, verificou-se diferença estatística entre as respostas médias do CONWIP e DBR.

O CONWIP apresentou controle mais efetivo do estoque em processo, resultando em menor tempo médio de fluxo. Isso teve o custo de deixar as ordens aguardando por mais tempo

no *pre-shop pool*, o que também contribuiu para maior variação dos resultados em relação ao DBR. Ainda assim, como não foram observadas diferenças estatísticas entre a média do *lead time* nos dois modelos, esses resultados sugerem que o controle da liberação de ordens pelo estoque físico com o CONWIP pode ser vantajoso em algumas situações.

Para avaliar a influência dos fatores no desempenho dos experimentos, além da análise descritiva dos resultados, realizou-se uma Análise de Variância. A ANOVA das variáveis de resposta, baseada nos experimentos hierárquico-cruzados, evidenciou que os fatores: modelo, severidade do gargalo e posição do gargalo influenciaram os indicadores tempo médio de fluxo, *lead time* médio, atraso médio e percentual de ordens atrasadas.

Para a variável utilização média do gargalo não foi feita a ANOVA, uma vez que essa apresentou variação mínima dos resultados e estava relacionada com a demanda (parâmetro da taxa de chegada) ao invés dos fatores dos experimentos. Além disso, como o limite de *buffer* estava condicionado ao fator modelo, não foi possível analisar o efeito principal do mesmo.

Quanto ao impacto das interações na ANOVA, se mostraram significativas para o tempo médio de fluxo: “Modelo x Posição”, “Modelo x Severidade”, “*Buffer* x Severidade” e “Posição x Severidade”. Para o *lead time* médio, foram estatisticamente significativas: “Modelo x Severidade” e “Posição x Severidade”. Para as respostas do atraso médio e percentual de atraso, evidenciou-se influência das interações: “Modelo x *Buffer*”, “Modelo x Posição”, “Modelo x Severidade”, “Posição x Severidade”.

Verificou-se que para todos os indicadores, a interação entre os fatores “*Buffer* x Posição” não se mostrou significativa. Uma das possíveis razões para essa evidência, é o ajuste do nível de proteção (*tbp*) de acordo com a posição do gargalo, a fim de representar na modelagem do DBR o dimensionamento do *time buffer* da restrição.

Como discutido na problematização, a avaliação do impacto da posição do gargalo tem ganhado evidência nos estudos sobre CONWIP e DBR e contribuído para identificação de diferenças entre os modelos. O presente estudo corroborou com pesquisas anteriores, evidenciando que no *assembly flow shop* em análise, o CONWIP não foi fortemente afetado pela mudança na posição do gargalo.

O CONWIP apresentou desempenho um pouco melhor quando o gargalo estava localizado nas primeiras estações (g=1 e g=3), devido a um controle mais efetivo do sistema. Observou-se, contudo, um comprometimento maior do desempenho do modelo, quando o gargalo estava mais próximo ao final do sistema (g=7 e g=9).

Já no DBR, verificou-se que para os indicadores *lead time* médio, atraso médio e percentual de atraso, os desempenhos foram inferiores nas extremidades do sistema (máquinas

1 e 9), enquanto que o melhor desempenho foi na posição 3 (última máquina da linha de fabricação). Quanto ao tempo médio de fluxo, os resultados do DBR indicaram que o desempenho do sistema piorou à medida que a posição do gargalo era mais próxima do final do sistema. Apesar das variações dos resultados, pode-se afirmar que os dois modelos apresentam comportamento semelhante quando o gargalo é a última máquina do sistema.

Em relação a construção dos modelos de simulação, observou-se que apesar das semelhanças entre os modelos, o DBR possui uma estrutura mais complexa de programação e controle. Mesmo com o foco no gargalo, o controle da liberação por meio do tempo, exige o monitoramento contínuo dos tempos de produção (sujeitos a variabilidades). Enquanto que o CONWIP, com o foco principal o controle do estoque (físico) em processo, realizou mais efetivamente essa atividade e é flexível para ser combinado com diversas regras de sequenciamento ou otimização da programação dos *jobs*.

De modo geral, o CONWIP apresenta menos requisitos para configuração e implantação, até mesmo para estudos teóricos de simulação. No ambiente de montagem analisado, os principais desafios remeteram a definição dos *loops* de controle e aos limites de WIP. Como optou-se por apenas dois segmentos, esses contemplaram desde as primeiras estações da fabricação até o processo de finalização. O modelo já considerava todo o sistema (segmentos críticos), inclusive a estação gargalo (independentemente da localização da mesma nos experimentos).

No DBR clássico, em contrapartida, as políticas são dependentes da correta identificação do gargalo e tornam-se defasadas quando ocorre a mudança do gargalo. É preciso saber não apenas a posição do gargalo, como também estimar as principais variabilidades do sistema para corretamente dimensionar os *time buffers* de expedição, montagem e do gargalo. Foram ajustados os parâmetros para a programação e o controle da liberação, de acordo com a posição do gargalo em cada conjunto de instâncias. No estudo, devido a natureza teórica, as estimativas foram decorrentes da observação do desempenho do CONWIP. Além disso, buscou-se simular o DBR o mais próximo do que preconiza a literatura, permitindo o cálculo do momento a liberação dos *jobs* de acordo com o desempenho do sistema e a proteção com os *time buffers*.

Diante do exposto, considera-se que o presente estudo cumpriu o objetivo pretendido e estendeu a discussão de trabalhos anteriores sobre a comparação dos modelos CONWIP e DBR em um ambiente com fluxo convergente. Assim como, demonstrou a importância de considerar fatores como a posição do gargalo e a severidade desse diante da implantação dos modelos em questão.

10.2 Limitações do trabalho e desdobramentos

Apesar das contribuições dos resultados e discussões apresentadas nas seções anteriores, a presente pesquisa possui limitações quanto a abrangência das conclusões. Um fator relevante, a ser considerado em pesquisas futuras, são as características do ambiente de produção. Como esta era uma das primeiras comparações do modelo em sistemas de montagem, optou-se pela análise de um ambiente de produção simples, onde todos os produtos seguiam o mesmo roteiro e havia simetria entre as linhas de fabricação (mesma quantidade de operações e tempos médios).

Observam-se oportunidades de pesquisas que contemplem ambientes como o *assembly job shop*. Espera-se que nesse ambiente, as diferenças entre os modelos se tornem mais evidentes, pois sistemas simples como *flow shops* tendem a favorecer o CONWIP. Assim como, seja mais complexa a aplicação e configuração dos modelos. Além disso, como os estudos sobre esses mecanismos são sobretudo teóricos, considera-se importante incrementar gradativamente a complexidade dos sistemas simulados, a fim de que esses representem de forma mais próxima os desafios de um sistema real, incluindo variabilidades como quebras de máquinas, etc.

Em relação a representação do DBR, considera-se que o presente estudo destacou-se em relação as pesquisas anteriores ao utilizar o conceito de *time buffer* ao invés de WIP, conforme preconiza a literatura clássica do DBR. No entanto, foram feitas simplificações quanto aos conceitos da TOC a fim de evitar a programação dos *jobs* com a consideração de capacidade finita do gargalo, uma vez que se desejava representar a dinâmica da produção MTO. Como também, omitiu-se o gerenciamento de pulmões, que torna o controle das ordens em processo mais efetivo, especialmente, em ambientes de produção com roteiros mais complexos.

Outro conceito relevante da TOC, sugerido para pesquisas futuras, é o S-DBR. A literatura sobre a versão simplificada do DBR ainda é escassa e os estudos comparativos do CONWIP com o DBR não fazem menção ao método simplificado. No S-DBR há apenas o pulmão de mercado e considera-se o mercado é a principal restrição do sistema, mesmo quando há restrição de capacidade. Além disso, não é feita a programação detalhada do recurso crítico.

Em relação ao CONWIP, as principais limitações do estudo foram a não consideração do ajuste dinâmico de cartões, que poderia trazer vantagens ao modelo, e a adoção de uma regra simples de sequenciamento da fila de entrada. Nesse sentido, observa-se a oportunidade de estudos que discutam as diferenças entre os modelos no ambiente de montagem com a adoção de regras mais complexas ou técnicas otimizantes.

REFERÊNCIAS

AGLAN, C.; DURMUSOGLU, M. B. Lot-splitting approach of a hybrid manufacturing system under CONWIP production control: a mathematical model. **International Journal of Production Research**, v. 53, n. 5, p. 1561–1583, 2015.

AJORLOU, S.; SHAMS, I. Artificial bee colony algorithm for CONWIP production control system in a multi-product multi-machine manufacturing. **Journal of Intelligent Manufacturing**, v. 24, p. 1145–1156, 2013.

AKINC, U.; MEREDITH, J. R. Make-to-forecast: Customization with fast delivery. **International Journal of Operations and Production Management**, v. 35, n. 5, p. 728–750, 2015.

ATAN, Z.; AHMADI, T.; STEGEHUIS, C.; KOK, T. de; ADAN, I. Assemble-to-Order Systems: A Review. **European Journal of Operational Research**, v. 261, n. 3, p. 866–879, 2017.

ATWATER, J. B.; CHAKRAVORTY, S. S. A study of the utilization of capacity constrained resources in Drum-Buffer-Rope systems. **Production and Operations Management**, v. 11, n. 2, p. 259–273, 2002.

BAHAJI, N.; KUHL, M. E. A simulation study of new multi-objective composite dispatching rules, CONWIP, and push lot release in semiconductor fabrication. **International Journal of Production Research**, v. 46, n. 14, p. 3801–3824, 2008.

BAHOUTH, S. B.; FOOTE, B. L. Managing a two-bottleneck job shop with a two-machine flow shop algorithm. **International Journal of Production Research**, v. 32, n. 10, p. 2463–2477, 1994.

BEEMSTERBOER, B.; LAND, M.; TEUNTER, R. Hybrid MTO-MTS production planning: An explorative study. **European Journal of Operational Research**, v. 248, n. 2, p. 453–461, 2016.

BEEMSTERBOER, B.; LAND, M.; TEUNTER, R.; BOKHORST, J. Integrating make-to-order and make-to-stock in job shop control. **International Journal of Production Economics**, v. 185, n. March 2017, p. 1–10, 2017.

BELISÁRIO, L. S.; PIERREVAL, H. Using genetic programming and simulation to learn how to dynamically adapt the number of cards in reactive pull systems. **Expert Systems with Applications**, v. 42, n. 6, p. 3129–3141, 2015.

BERGAMASCHI, D.; CIGOLINI, R.; PERONA, M.; PORTIOLI, A. Order review and release strategies in a job shop environment: A review and a classification. **International Journal of Production Research**, v. 35, n. 2, p. 399–420, 1997.

BERTRAND, J. W. M.; FRANSOO, J. C. Operations management research methodologies using quantitative modeling. **International Journal of Operations & Production Management**, v. 22, n. 2, p. 241–264, 2002.

BERTRAND, J. W. M.; VAN DE WAKKER, A. M. An investigation of order release and flow time allowance policies for assembly job shops. **Production Planning and Control**, v. 13, n. 7, p. 639–648, 2002.

BETTERTON, C. E.; COX, J. F. Espoused drum-buffer-rope flow control in serial lines: A comparative study of simulation models. **International Journal of Production Economics**,

v. 117, n. 1, p. 66–79, 2009.

BETTERTON, C. E.; SILVER, S. J. Detecting bottlenecks in serial production lines - A focus on interdeparture time variance. **International Journal of Production Research**, v. 50, n. 15, p. 4158–4174, 2012.

BLACKSTONE JR, J. H. Revisão da literatura sobre tambor-pulmão-corda, gerenciamento de pulmões e distribuição. In: COX III, J. F.; SCHLEIER JR, J. G. (Ed.). **Handbook da Teoria das Restrições**. 1. ed. Porto Alegre: Bookman, 2013. p. 151–178.

BLACKSTONE JR, J. H.; COX III, J. F. Designing unbalanced lines - understanding protective capacity and protective inventory. **Production Planning & Control**, v. 13, n. 4, p. 416–423, 2002.

BRAGLIA, M.; FROSOLINI, M.; GABBRIELLI, R.; ZAMMORI, F. CONWIP card setting in a flow-shop system with a batch production machine. **International Journal of Industrial Engineering Computations**, v. 2, n. 1, p. 1–18, 2011.

CAO, D.; CHEN, M. A mixed integer programming model for a two line CONWIP-based production and assembly system. **International Journal of Production Economics**, v. 95, n. 3, p. 317–326, 2005.

CHAKRAVORTY, S. S.; ATWATER, J. B. The impact of free goods on the performance of drum-buffer-rope scheduling systems. **International Journal of Production Economics**, v. 95, n. 3, p. 347–357, 2005.

CHEN, J.; HUANG, G. Q.; LUO, H.; WANG, J. Synchronisation of production scheduling and shipment in an assembly flowshop. **International Journal of Production Research**, v. 53, n. 9, p. 2787–2802, 2015.

CHINET, F. S.; GODINHO FILHO, M. Sistema POLCA: revisão, classificação e análise da literatura. **Gestão & Produção**, v. 21, n. 3, p. 532–542, 2014.

CONWAY, R.; MAXWELL, W.; MCCLAIN, J. O.; THOMAS, L. J. The Role of Work-in-Process Inventory in Serial Production Lines. **Operations Research**, v. 36, n. 2, p. 229–241, 1988.

CORBETT, T.; CSILLAG, J. M. Analysis of the effects of seven Drum-Buffer-Rope implementations. **Production and Inventory Management Journal**, v. 42, n. 3, 2001.

COX III, J. F.; SCHLEIER JR, J. G. **Theory of Constraints Handbook**. 1. ed. [s.l.] McGraw-Hill Education, 2010.

COX III, J. F.; SPENCER, M. S. **Manual da Teoria das Restrições**. Porto Alegre: Bookman, 2002.

CRAIGHEAD, C. W.; PATTERSON, J. W.; FREDENDALL, L. D. Protective capacity positioning: Impact on manufacturing cell performance. **European Journal of Operational Research**, v. 134, n. 2, p. 425–438, 2001.

DANIEL, C. R. A. **Estudo do desempenho de métodos de Lenth na análise de planos do tipo Split-Plot na ausência de réplicas**. 2008. 82f. Dissertação (Mestrado em Estatística) - Universidade Federal de Pernambuco, Recife, 2008.

DUENYAS, I. Estimating the throughput of a cyclic assembly system. **International Journal of Production Research**, v. 32, n. 6, p. 1403–1419, 1994a.

DUENYAS, I. A simple release policy for networks of queues with controllable inputs. **Operations Research**, v. 42, n. 6, p. 1162–1171, 1994b.

DUENYAS, I.; HOPP, W. J.; SPEARMAN, M. L. Characterizing the Output Process of a CONWIP Line with Deterministic Processing and Random Outages. **Management Science**, v. 39, n. 8, p. 975–988, 1993.

DURI, C.; FREIN, Y.; LEE, H. S. Performance evaluation and design of a CONWIP system with inspections. **International Journal of Production Economics**, v. 64, n. 1, p. 219–229, 2000.

DURMUSOGLU, M. B.; AGLAN, C. CONWIP and Hybrid CONWIP Production Control Systems: A Literature Review. In: KHOJASTEH, Y. (Ed.). **Production Management: Advanced Models, Tools, and Applications for Pull Systems**. Boca Raton: CRC Press, 2018. p. 123–142.

EBADIAN, M.; RABBANI, M.; TORABI, S. A.; JOLAI, F. Hierarchical production planning and scheduling in make-to-order environments: reaching short and reliable delivery dates. **International Journal of Production Research**, v. 47, n. 20, p. 5761–5789, 2009.

ELFTMAN, T. M. **Examination of the effects of bottlenecks and production control rules at assembly stations**. 1999. University of Florida, 1999.

FERNANDES, F. C. F.; GODINHO FILHO, M. Production control systems: Literature review, classification, and insights regarding practical application. **African Journal of Business Management**, v. 5, n. 14, p. 5573–5582, 2011.

FRAMINAN, J. M.; GONZÁLEZ, P. L.; RUIZ-USANO, R. The CONWIP production control system: Review and research issues. **Production Planning & Control**, v. 14, n. 3, p. 255–265, 2003.

FRAMINAN, J. M.; GONZÁLEZ, P. L.; RUIZ-USANO, R. Dynamic card controlling in a Conwip system. **International Journal of Production Economics**, v. 99, p. 102–116, 2006.

FRAMINAN, J. M.; RUIZ-USANO, R.; LEISTEN, R. Input control and dispatching rules in a dynamic CONWIP flow-shop. **International Journal of Production Research**, v. 38, n. 18, p. 4589–4598, 2000.

FRAMINAN, J. M.; RUIZ-USANO, R.; LEISTEN, R. Sequencing CONWIP flow-shops: Analysis and heuristics. **International Journal of Production Research**, v. 39, n. 12, p. 2735–2749, 2001.

FREDENDALL, L. D.; LEA, B. R. Improving the product mix heuristic in the theory of constraints. **International Journal of Production Research**, v. 35, n. 6, p. 1535–1544, 1997.

FREDENDALL, L. D.; OJHA, D.; PATTERSON, J. W. Concerning the theory of workload control. **European Journal of Operational Research**, v. 201, n. 1, p. 99–111, 2010.

FRY, T. D.; COX, J. F.; BLACKSTONE JR., J. H. An analysis and discussion of the Optimized Production Technology software and its use. **Production and Operations Management**, v. 1, n. 2, p. 229–242, 1992.

FRY, T. D.; KARWAN, K. R.; STEELE, D. C. Implementing Drum-Buffer-Rope to Control Manufacturing Lead Time. **The International Journal of Logistics Management**, v. 2, n. 1, p. 12–18, 1991.

FRY, T. D.; OLIFF, M. D.; MINOR, E. D.; KEONG LEONG, G. The effects of product structure and sequencing rule on assembly shop performance. **International Journal of Production Research**, v. 27, n. 4, p. 671–686, 1989.

FRY, T. D.; PHILIPOOM, P. R.; LEONG, G. K.; SMITH, A. E. An Investigation of Bottleneck Position in a Multi-Stage Job Shop. **International Journal of Operations &**

Production Management, v. 7, n. 6, p. 55–63, 1987.

GARMDARE, H. S.; LOTFI, M. M.; HONARVAR, M. Integrated model for pricing, delivery time setting, and scheduling in make-to-order environments. **Journal of Industrial Engineering International**, v. 14, n. 1, p. 55–64, 2018.

GILLAND, W. G. A simulation study comparing performance of CONWIP and bottleneck-based release rules. **Production Planning & Control**, v. 13, n. 2, p. 211–219, 2002.

GOLDRATT, E. M. **The Haystack Syndrome: Shifting Information Out of the Data Ocean**. [s.l.] North River Press, 1990.

GOLMOHAMMADI, D. A study of scheduling under the theory of constraints. **International Journal of Production Economics**, v. 165, p. 38–50, 2015.

GRABENSTETTER, D. H.; USHER, J. M. Developing due dates in an engineer-to-order engineering environment. **International Journal of Production Research**, v. 7543, n. March, p. 1–13, 2017.

GUIDE JR., V. D. R.; GHISELLI, G. A. Implementation of drum-buffer-rope at a military rework depot engine works. **Production and Inventory Management Journal**, v. 36, n. 3, p. 79, 1995.

HADAS, L.; CYPLIK, P.; FERTSCH, M. Method of buffering critical resources in make-to-order shop floor control in manufacturing complex products. **International Journal of Production Research**, v. 47, n. 8, p. 2125–2139, 2009.

HARROD, S.; KANET, J. J. Applying work flow control in make-to-order job shops. **International Journal of Production Economics**, v. 143, n. 2, p. 620–626, 2013.

HENDRY, L. C.; KINGSMAN, B. G. Production planning systems and their applicability to make-to-order companies. **European Journal of Operational Research**, v. 40, p. 1–15, 1989.

HENDRY, L.; HUANG, Y.; STEVENSON, M. Workload control: Successful implementation taking a contingency-based view of production planning and control. **International Journal of Operations & Production Management**, v. 33, n. 1, p. 69–103, 2013.

HERER, Y. T.; MASIN, M. Mathematical programming formulation of CONWIP based production lines; and relationships to MRP. **International Journal of Production Research**, v. 35, n. 4, p. 1067–1076, 1997.

HOPP, W. J.; ROOF, M. L. Setting WIP levels with statistical throughput control (STC) in CONWIP production lines. **International Journal of Production Research**, v. 36, n. 4, p. 867–882, 1998.

HOPP, W. J.; SPEARMAN, M. L. **A Ciência da Fábrica**. 3. ed. Porto Alegre: Bookman, 2013.

HUANG, G.; CHEN, J.; WANG, X.; SHI, Y. A simulation study of CONWIP assembly with multi-loop in mass production, multi-products and low volume and OKP environments. **International Journal of Production Research**, v. 53, n. 14, p. 4160–4175, 2015.

HUANG, M.; WANG, D.; IP, W. H. Simulation study of CONWIP for a cold rolling plant. **International Journal of Production Economics**, v. 54, n. 3, p. 257–266, 1998.

IKEZIRI, L. M.; SOUZA, F. B.; GUPTA, M. C.; FIORINI, P. C. Theory of constraints: review and bibliometric analysis. **International Journal of Production Research**, p. 1–35,

2018.

IOANNIDIS, S.; KOUIKOGLOU, V. S. Revenue management in single-stage CONWIP production systems. **International Journal of Production Research**, v. 46, n. 22, p. 6513–6532, 2008.

JAEGLER, Y.; JAEGLER, A.; BURLAT, P.; LAMOURI, S.; TRENTESAUX, D. The ConWip production control system: a systematic review and classification. **International Journal of Production Research**, v. 7543, n. September, p. 1–22, 2017.

JODLBAUER, H.; HUBER, A. Service-level performance of MRP, kanban, CONWIP and DBR due to parameter stability and environmental robustness. **International Journal of Production Research**, v. 46, n. 8, p. 2179–2195, 2008.

KHOJASTEH-GHAMARI, Y. A performance comparison between Kanban and CONWIP controlled assembly systems. **Journal of Intelligent Manufacturing**, v. 20, n. 6, p. 751–760, 2009.

KIM, S.; COX, J. F.; MABIN, V. J. An exploratory study of protective inventory in a re-entrant line with protective capacity. **International Journal of Production Research**, v. 48, n. 14, p. 4153–4178, 2010.

KIM, S.; DAVIS, K. R.; COX III, J. F. An investigation of output flow control, bottleneck flow control and dynamic flow control mechanisms in various simple lines scenarios. **Production Planning and Control**, v. 14, n. 1, p. 15–32, 2003.

KINGSMAN, B. G.; TATSIOPOULOS, I. P.; HENDRY, L. C. A structural methodology for managing manufacturing lead times in make-to-order companies. **European Journal of Operational Research**, v. 40, n. 2, p. 196–209, 1989.

KOH, S.; BULFIN, R. L. Comparison of DBR with CONWIP in an unbalanced production line with three stations. **International Journal of Production Research**, v. 42, n. 2, p. 391–404, 2004.

LAGE JUNIOR, M.; GODINHO FILHO, M. Variations of the kanban system: Literature review and classification. **International Journal of Production Economics**, v. 125, n. 1, p. 13–21, 2010.

LAW, A. M.; KELTON, W. D. **Simulation Modeling and Analysis**. 3. ed. Boston: McGraw-Hill, 2000.

LAWRENCE, S. R.; BUSS, A. H. Shifting production bottlenecks: causes, cures, and conundrums. **Production and Operations Management**, v. 3, n. 1, p. 21–37, 1994.

LEE, H.; SEO, D. W. Performance evaluation of WIP-controlled line production systems with constant processing times. **Computers and Industrial Engineering**, v. 94, p. 138–146, 2016.

LEE, J. H.; CHANG, J. G.; TSAI, C. H.; LI, R. K. Research on enhancement of TOC Simplified Drum-Buffer-Rope system using novel generic procedures. **Expert Systems with Applications**, v. 37, n. 5, p. 3747–3754, 2010.

LEU, B. Generating a backlog list for a CONWIP production line : A simulation study. **Production Planning & Control**, v. 11, n. 4, p. 409–418, 2000.

LIU, Q.; HUANG, D. Dynamic card number adjusting strategy in card-based production system. **International Journal of Production Research**, v. 47, n. 21, p. 6037–6050, 2009.

LOUW, L.; PAGE, D. C. Queuing network analysis approach for estimating the sizes of the time buffers in Theory of Constraints-controlled production systems. **International Journal**

of Production Research, v. 42, n. 6, p. 1207–1226, 2004.

LU, H. L.; HUANG, G. Q.; YANG, H. D. Integrating order review/release and dispatching rules for assembly job shop scheduling using a simulation approach. **International Journal of Production Research**, v. 49, n. 3, p. 647–669, 2011.

MABIN, V. J.; BALDERSTONE, S. J. The performance of the theory of constraints methodology. **International Journal of Operations & Production Management**, v. 23, n. 6, p. 568–595, 2003.

MIGUEL, P. A. C. Estudo de caso na engenharia de produção: estruturação e recomendações para sua condução. **Produção**, v. 17, n. 1, p. 216–229, 2007.

MILLSTEIN, M. A.; MARTINICH, J. S. Takt Time Grouping: implementing kanban-flow manufacturing in an unbalanced, high variation cycle-time process with moving constraints. **International Journal of Production Research**, v. 52, n. 23, p. 6863–6877, 2014.

MONTGOMERY, D. C. **Design and Analysis of Experiments**. 9. ed. [s.l.] John Wiley & Sons, 2017.

MORABITO NETO, R.; PUREZA, V. Modelagem e Simulação. In: MIGUEL, P. A. C. (Ed.). **Metodologia de pesquisa em Engenharia de Produção e Gestão de Operações**. 2. ed. Rio de Janeiro: Elsevier, 2012. p. 169–198.

NG, A. H. C.; BERNEDIXEN, J.; SYBERFELDT, A. A comparative study of production control mechanisms using simulation-based multi-objective optimisation. **International Journal of Production Research**, v. 50, n. 2, p. 359–377, 2012.

OHNO, T. **Toyota Production System: Beyond Large-Scale ProductionProductivity Press**, 1988. .

OLHAGER, J. Strategic positioning of the order penetration point. **International Journal of Production Economics**, v. 85, n. 3, p. 319–329, 2003.

OLHAGER, J. The role of the customer order decoupling point in production and supply chain management. **Computers in Industry**, v. 61, n. 9, p. 863–868, 2010.

PEREIRA, M. T.; SANTORO, M. C. An integrative heuristic method for detailed operations scheduling in assembly job shop systems. **International Journal of Production Research**, v. 49, n. 20, p. 6089–6105, 2011.

PINEDO, M. L. **Scheduling: Theory, Algorithms, and Systems**. 4. ed. [s.l.] Springer, 2012.

PRAKASH, J.; CHIN, J. F. Modified CONWIP systems: a review and classification. **Production Planning & Control**, v. 26, n. 4, p. 296–307, 2015.

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

RADOVILSKY, Z. D. A quantitative approach to estimate the size of the time buffer in the theory of constraints. **International Journal of Production Economics**, v. 55, n. 2, p. 113–119, 1998.

RAHMAN, S. Theory of constraints: A review of the philosophy and its applications. **International Journal of Operations & Production Management**, v. 18, n. 4, p. 336–355, 1998.

ROCHA, F.; SILVA, E.; LOPES, Â.; DIAS, L.; PEREIRA, G.; FERNANDES, N. O.; CARMO-SILVA, S. Materials Flow Control in Hybrid Make-to-Stock/Make-to-Order Manufacturing. In: CORMAN, F.; VOSS, S.; NEGENBORN, R. (Ed.). **Computational**

Logistics. ICCL 2015. Lecture Notes in Computer Science. [s.l.] Springer, Cham, 2015. 9335p. 559–568.

RONEN, B.; STARR, M. K. Synchronized Manufacturing as in OPT: from practice to theory. **Computers and Industrial Engineering**, v. 18, n. 4, p. 585–600, 1990.

RUSSELL, G. R.; FRY, T. D. Order review/release and lot splitting in drum-buffer-rope. **International Journal of Production Research**, v. 35, n. 3, p. 827–845, 1997.

RYAN, S. M.; BAYNAT, B.; CHOOBINEH, F. Determining inventory levels in a CONWIP controlled job shop. **IIE Transactions (Institute of Industrial Engineers)**, v. 32, n. 2, p. 105–114, 2000.

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

SCHRAGENHEIM, E.; RONEN, B. Drum-buffer-rope shop floor control. **Production and Inventory Management Journal**, v. 31, n. 3, p. 18–22, 1990.

SCHRAGENHEIM, E.; RONEN, B. Buffer Managment - a Diagnostic Tool for Production Control. **Production and Inventory Management Journal**, v. Second Qua, p. 74–79, 1991.

SIMONS JR., J. V.; SIMPSON III, W. P. An Exposition of Multiple Constraint Scheduling As Implemented in the Goal System (Formerly Disaster). **Production and Operations Management**, v. 6, n. 1, p. 3–22, 1997.

SIRIKRAI, V.; YENRADEE, P. Modified drum-buffer-rope scheduling mechanism for a non-identical parallel machine flow shop with processing-time variation. **International Journal of Production Research**, v. 44, n. 17, p. 3509–3531, 2006.

SOEPENBERG, G. D.; LAND, M.; GAALMAN, G. The order progress diagram: A supportive tool for diagnosing delivery reliability performance in make-to-order companies. **International Journal of Production Economics**, v. 112, n. 1, p. 495–503, 2008.

SOUZA, F. B. Do OPT à Teoria das Restrições: avanços e mitos. **Produção**, v. 15, n. 2, p. 184–197, 2005.

SOUZA, F. B.; BAPTISTA, H. R. Proposta de avanço para o método Tambor-Pulmão-Corda Simplificado aplicado em ambientes de produção sob encomenda. **Gestão & Produção**, v. 17, n. 4, p. 735–746, 2010.

SPEARMAN, M.; WOODRUFF, D.; HOPP, W. CONWIP: a pull alternative to kanban. **International Journal of Production Research**, v. 28, n. 5, p. 879–894, 1990.

SPENCER, M. S.; COX, J. F. Optimum production technology (OPT) and the theory of constraints (TOC): analysis and genealogy. **International Journal of Production Research**, v. 33, n. 6, p. 1495–1504, 1995.

SRIKANTH, M. (Shri). TPC, gerenciamento de pulmões e classificação de fluxo VATI. In: COX III, J. F.; SCHLEIER JR, J. G. (Ed.). **Handbook da Teoria das Restrições**. 1. ed. Porto Alegre: Bookman, 2013. p. 1206.

STEVENSON, M.; HENDRY, L. C.; KINGSMAN, B. G. A review of production planning and control: the applicability of key concepts to the make-to-order industry. **International Journal of Production Research**, v. 43, n. 5, p. 869–898, 2005.

STEWART, R. **Simulation: the practice of model development and use**. [s.l: s.n.]

THÜRER, M.; FERNANDES, N. O.; STEVENSON, M.; QU, T. On the backlog-sequencing decision for extending the applicability of ConWIP to high-variety contexts: an assessment by

simulation. **International Journal of Production Research**, v. 7543, n. April, p. 1–17, 2017a.

THÜRER, M.; QU, T.; STEVENSON, M.; LI, C. D.; HUANG, G. Q. Deconstructing bottleneck shiftiness: the impact of bottleneck position on order release control in pure flow shops. **Production Planning and Control**, v. 28, n. 15, p. 1223–1235, 2017b.

THÜRER, M.; STEVENSON, M. On the beat of the drum: improving the flow shop performance of the Drum–Buffer–Rope scheduling mechanism. **International Journal of Production Research**, v. 7543, n. November, p. 1–12, 2017.

THÜRER, M.; STEVENSON, M. Bottleneck-oriented order release with shifting bottlenecks: An assessment by simulation. **International Journal of Production Economics**, v. 197, n. January, p. 275–282, 2018.

THÜRER, M.; STEVENSON, M.; PROTZMAN, C. W. Card-based production control: a review of the control mechanisms underpinning Kanban, ConWIP, POLCA and COBACABANA systems. **Production Planning & Control**, v. 27, n. 14, p. 1143–1157, 2016.

THÜRER, M.; STEVENSON, M.; SILVA, C. Three decades of workload control research: a systematic review of the literature. **International Journal of Production Research**, v. 49, n. 23, p. 6905–6935, 2011.

THÜRER, M.; STEVENSON, M.; SILVA, C.; HUANG, G. The application of workload control in assembly job shops: An assessment by simulation. **International Journal of Production Research**, v. 50, n. 18, p. 5048–5062, 2012.

THÜRER, M.; STEVENSON, M.; SILVA, C.; QU, T. Drum-buffer-rope and workload control in High-variety flow and job shops with bottlenecks: An assessment by simulation. **International Journal of Production Economics**, v. 188, p. 116–127, 2017c.

TRIOLA, M. F. **Introdução à estatística**. 9. ed. [s.l.] LTC, 2005.

TU, Y. M.; LI, R. K. Constraint time buffer determination model. **International Journal of Production Research**, v. 36, n. 4, p. 1091–1103, 1998.

UMBLE, M. M.; SRIKANTH, M. L. **Synchronous manufacturing: Principles for world class excellence**. Cincinnati: South-Western Publishing Company, 1990.

WATSON, K. J.; BLACKSTONE, J. H.; GARDINER, S. C. The evolution of a management philosophy: The theory of constraints. **Journal of Operations Management**, v. 25, n. 2, p. 387–402, 2007.

WEMMERLÖV, U. Assemble-To-Order Manufacturing: Implications for Materials Management. **Journal of Operations Management**, v. 4, n. 4, p. 347–368, 1984.

WINSTON, W. L. **Operations Research: Applications and Algorithms**. 4. ed. Austrália: Thomson Brooks/Cole, 2004.

YE, T.; HAN, W. Determination of buffer sizes for drum-buffer-rope (DBR)-controlled production systems. **International Journal of Production Research**, v. 46, n. 10, p. 2827–2844, 2008.

APÊNDICE A – BLOCOS DO *ANYLOGIC* EMPREGADOS NA SIMULAÇÃO

Para elaboração dos modelos de simulação no *software AnyLogic* versão *Professional 8.3.3*, utilizou-se basicamente oito blocos para representar o processo produtivo, expostos na Figura 56 e descritos a seguir:

a) *Source*: elemento empregado para a geração das entidades, representa a entrada dos *jobs* no sistema. As entidades são geradas conforme os parâmetros de taxa de chegada ou intervalo entre chegadas, etc. Nos modelos desta pesquisa foram chamados de "*chegada*".

b) *Queue*: bloco que representa uma fila. Utilizado, principalmente, para representar o estoque entre operações. Pode possuir capacidade limitada ou infinita. Permite a ordenação das *jobs*, conforme algum critério de prioridade. A ordenação padrão é *first-in, first-out*. Nos modelos desenvolvidos, esse elemento foi usado para representar as filas do *pre-shop pool* (denominadas de "*preshop1*" e "*preshop2*") e as filas entre as máquinas do sistema de montagem (chamadas de "fila" acompanhadas do índice da máquina que precediam).

c) *Delay*: bloco utilizado para representar operações. Os tempos de processamento podem ser determinísticos ou estocásticos, definidos conforme distribuição de probabilidade. Podem possuir capacidade superior a 1 quando deseja-se modelar o processamento de mais de uma entidade conjuntamente. Nos modelos de simulação desta pesquisa, foram denominados de acordo com o índice da máquina, "*m1*", "*m2*", ... , "*m9*", com capacidade de 1 *job*.

d) *Sink*: representa a saída do sistema ou a "destruição" da entidade. Permite a coleta final de dados sobre as entidades, como o *lead time*, tempo de fluxo, etc. Nos modelos foram chamados de "*saida*".

e) *Split*: bloco que permite a "multiplicação" de uma entidade. Para cada entidade original que entra nesse bloco, pode-se fazer várias cópias com as mesmas características do agente original ou completamente novas. Esse bloco não gasta tempo no modelo. Nos modelos CONWIP e DBR, esse elemento foi utilizado para gerar as ordens de fabricação. Então, com base nas características dos *jobs* que entravam no sistema, foram geradas duas ordens de fabricação com as mesmas

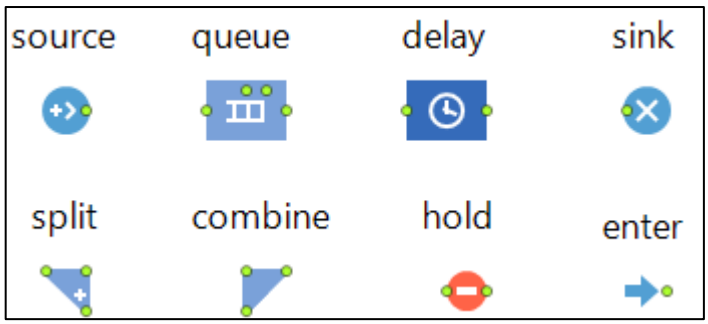
características (atributos) e numeração da entidade original. Utilizou-se a denominação "*copia*".

f) *Combine*: bloco utilizado para "unir" entidades. Com a entrada de duas entidades nesse bloco (uma por cada porta – círculos em verde na Figura 56), gera-se uma nova entidade, que pode ser criada com base nas características de uma das entidades originais ou completamente nova. Serve como um ponto de sincronização e também não consome tempo no modelo. Na modelagem dos sistemas, esse elemento foi aplicado para unir as ordens de fabricação, gerando uma entidade com os mesmos atributos (numeração, data prometida, etc.) para processamento na montagem. Foi denominado de "*comb*".

g) *Hold*: bloco que permite o controle de passagem das entidades. Utilizado para controlar o fluxo de produção no sistema de acordo com alguma condição especificada, que pode ser de tempo ou de carga de trabalho. Nos modelos CONWIP e DBR, foi utilizado para o controle da liberação de ordens para as linhas de fabricação. É um elemento que não consome tempo no modelo e foi chamado de "*control1*/2".

h) *Enter*: elemento que permite a transferência de agentes e a criação de fluxos personalizados no processo produtivo. Foi utilizado para pegar as ordens com a mesma numeração, que estavam aguardando na fila após a fabricação dos componentes, e levá-las para operação de montagem (artifício a fim de garantir a sincronização dessa operação).

Figura 56 - Blocos do *AnyLogic* empregados nos modelos de simulação

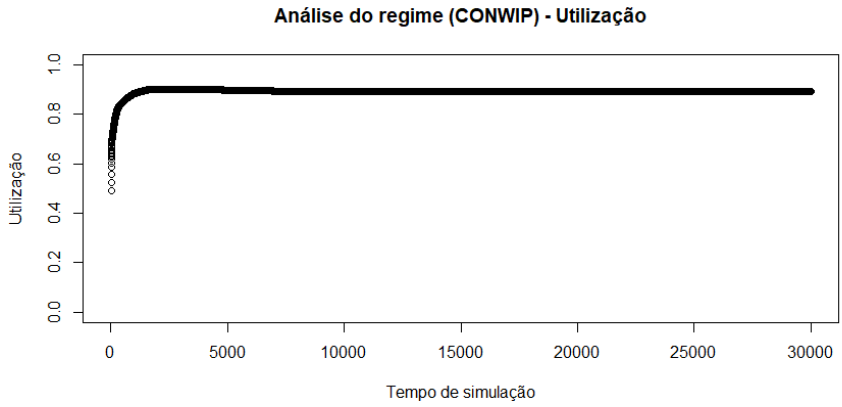


Fonte: Elaboração própria.

APÊNDICE B – ANÁLISE DOS REGIMES TRANSITÓRIO E PERMANENTE DAS VARIÁVEIS

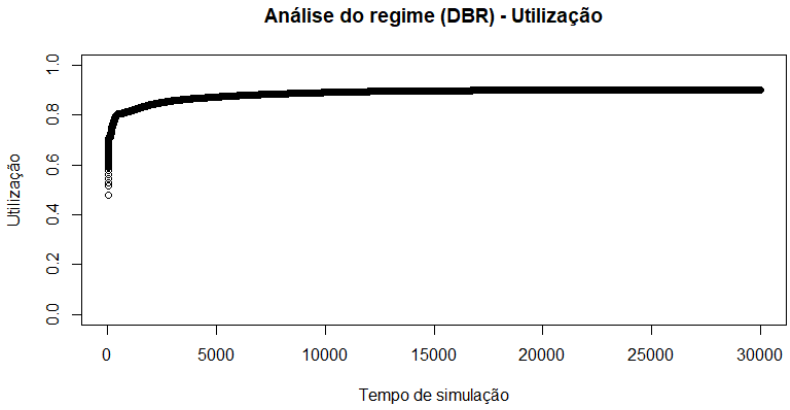
O comportamento de cada uma das variáveis de resposta dos experimentos é analisado por meio da observação das médias das variáveis (exceto o percentual) ao longo do tempo de simulação. As Figuras 57 e 58 mostram os resultados para a utilização do gargalo no cenário do CONWIP e DBR, respectivamente. Como pode-se observar, essa variável entrou no estado estacionário rapidamente nos dois casos, antes das primeiras 1.000 unidades de tempo da simulação.

Figura 57 - Análise do regime transitório para a variável utilização (CONWIP)



Fonte: Elaboração própria.

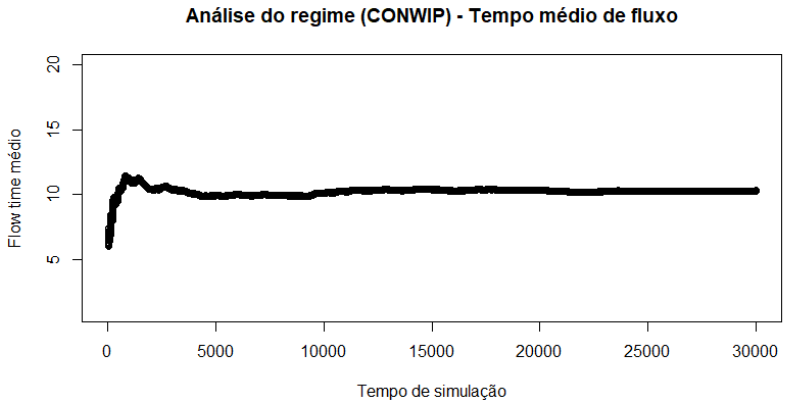
Figura 58 - Análise do regime transitório para a variável utilização (DBR)



Fonte: Elaboração própria.

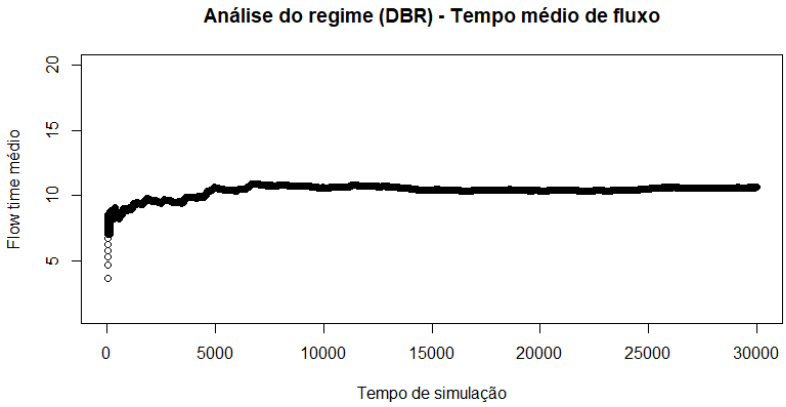
As Figura 59 e 60 expõem os resultados para o tempo médio de fluxo. A variável apresenta comportamento semelhante nos dois modelos ao longo da simulação. O tempo médio de fluxo entrou no regime estacionário próximo das 3.000 unidades de tempo da simulação, apresentando pequenas variações em torno do valor médio de 10 unidades.

Figura 59 - Análise do regime transitório para a variável tempo de fluxo (CONWIP)



Fonte: Elaboração própria.

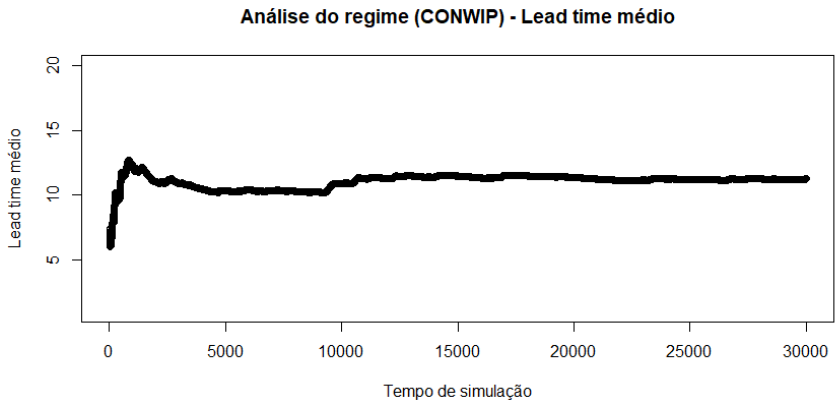
Figura 60 - Análise do regime transitório para a variável tempo de fluxo (DBR)



Fonte: Elaboração própria.

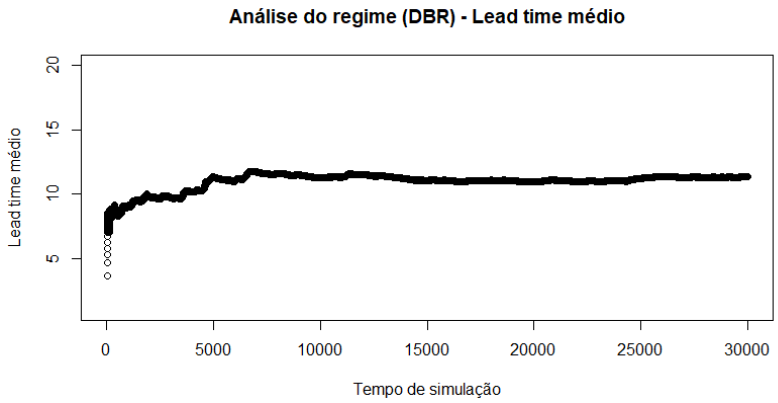
Em relação ao *lead time* médio, as Figuras 61 e 62 apresentam os resultados para o CONWIP e DBR, respectivamente. No caso do CONWIP, a variável entra em regime estacionário logo após as primeiras 2.000 unidades de tempo da simulação. Já no DBR, as oscilações se prolongaram por um pouco mais de tempo da simulação, mas a média da variável permaneceu entre 10 e 11 unidades de tempo.

Figura 61 - Análise do regime transitório para a variável *lead time* médio



Fonte: Elaboração própria.

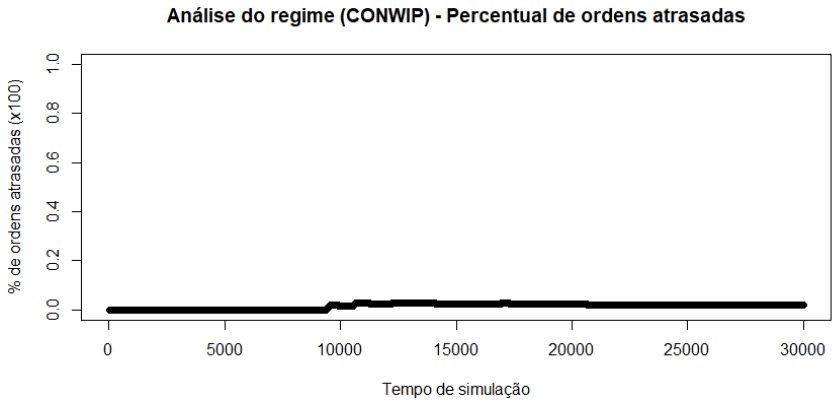
Figura 62 - Análise do regime transitório para a variável *lead time* médio (DBR)



Fonte: Elaboração própria.

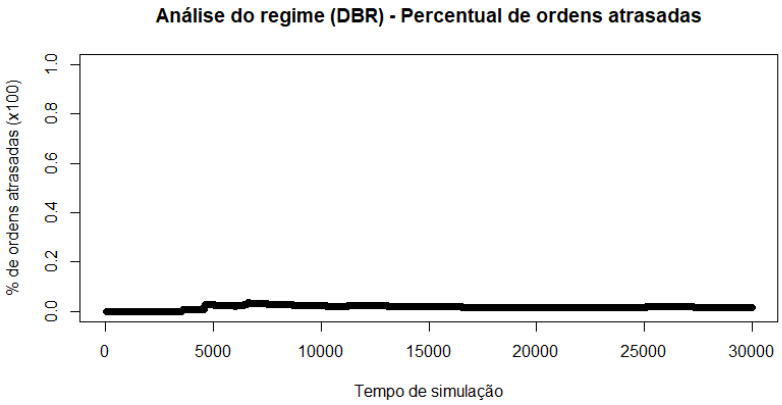
As Figuras 63 e 64 expõem os resultados para o percentual de ordens atrasadas nos modelos. Nos dois modelos, a variável apresentou variação em torno de valores bem baixos (menos de 5%), passando rapidamente do regime transitório ao permanente.

Figura 63 - Análise do regime transitório para a variável percentual de ordens atrasadas (CONWIP)



Fonte: Elaboração própria.

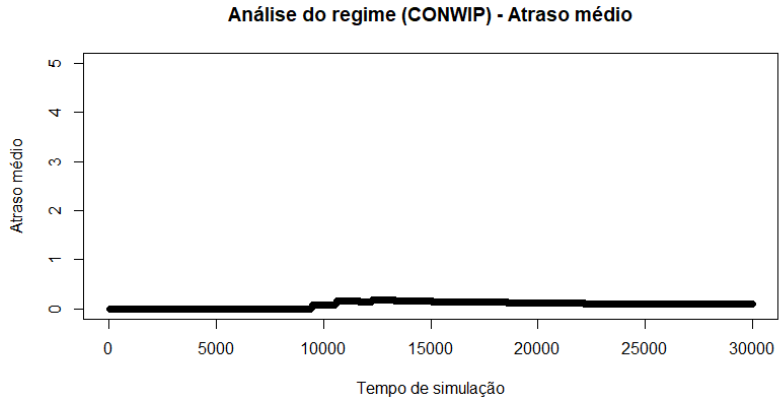
Figura 64 - Análise do regime transitório para a variável percentual de ordens atrasadas (DBR)



Fonte: Elaboração própria.

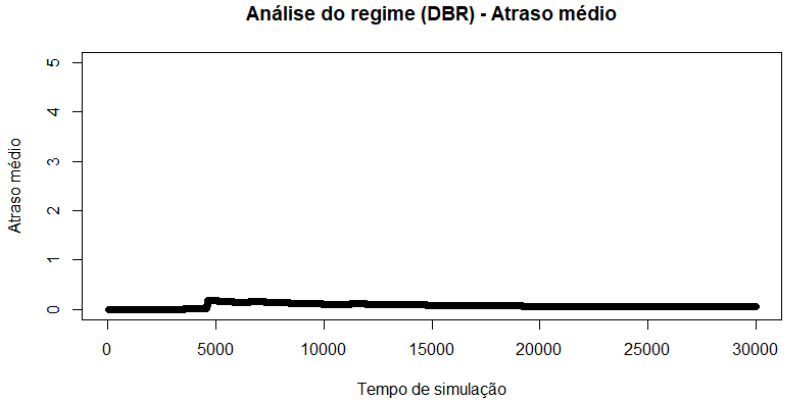
A variável de resposta atraso médio exibiu comportamento semelhante ao percentual de atraso. Em ambos os modelos, verifica-se comportamento estável desde os primeiros instantes da simulação. Como mostram as Figuras 65 e 66, os modelos apresentaram atraso médio muito baixo com pouquíssimas variações.

Figura 65 - Análise do regime transitório para a variável atraso médio (CONWIP)



Fonte: Elaboração própria.

Figura 66 - Análise do regime transitório para a variável atraso médio (DBR)



Fonte: Elaboração própria.