

APLICAÇÃO DE TEORIA DAS FILAS EM UMA AGÊNCIA BANCÁRIA DA CIDADE DE FARROUPILHA - RS

Bruna Caroline Orlandin (Universidade de Caxias do Sul - UCS) bcorlandin@ucs.br

Nathalia Tessari Moraes (Universidade de Caxias do Sul - UCS) ntmoraes1@ucs.br

Bruna Gabriele de Matos (Universidade de Caxias do Sul - UCS) bgmatos@ucs.br

Rafaela Boeira Cechin (Universidade de Caxias do Sul - UCS) rbcechin@ucs.br

Leandro Luís Corso (Universidade de Caxias do Sul - UCS) llcorso@ucs.br

Resumo

Em um mundo onde tudo acontece tão rápido, ficar alguns minutos parado em uma fila esperando por atendimento gera intolerância por parte dos usuários. Em agências bancárias, onde a competição para obter clientes é acirrada, faz-se necessário realizar um atendimento ágil e com qualidade visando conquistar os clientes. O objetivo deste trabalho é realizar um estudo que dimensione o sistema de atendimento de uma agência bancária situada na cidade de Farroupilha - RS, quanto ao seu atendimento, avaliando do ponto de vista operacional e propondo um sistema alterado. Para esta pesquisa realizou-se o levantamento dos dados e foi desenvolvido um modelo de otimização por meio de Teoria das filas. O estudo foi realizado durante o período de duas semanas e os resultados demonstram que o sistema atual da agência bancária é satisfatório, uma vez que, apresenta uma taxa de ocupação de 87%, o que justifica a fila ser observada e analisada em alguns períodos. Dessa forma, realizou-se a simulação de um cenário considerando quatro atendentes. A taxa de ocupação foi de 65%. Com quatro atendentes o sistema passaria a ser mais ágil, porém cabe ao gestor do sistema avaliar este cenário, pois com um atendente a mais a agência terá custos maiores.

Palavras-Chaves: Engenharia de Produção, Pesquisa Operacional, Teoria das Filas, Agência Bancária.

1. Introdução

Os serviços bancários se encaixam no que Albrecht (1992) ressalta quando diz que na filosofia de administração de serviços, toda organização deve atuar sendo um grande departamento de atendimento ao cliente. E para que o serviço prestado seja de qualidade, é necessário compreender a demanda de clientes, organizar sistemas que a controlem e ajustar a capacidade de atendimento em função da demanda, e como consequência, reduzir a espera e manter um serviço adequado ao cliente (GIANESE; CORRÊA, 1996).

No entanto, há diversas características que condicionam a operação de um sistema, estas podem ser classificadas em: forma dos atendimentos, forma das chegadas, disciplina da fila e estrutura do sistema (ANDRADE, 2009).

De acordo com Ricardo (2010), quando trata-se de serviços é mais difícil mensurar o grau de satisfação dos clientes, pois o que determina a qualidade é o valor percebido pelo cliente. No entanto, são os fatores intangíveis que são percebidos através dos sentidos, como, por exemplo, um ambiente acolhedor, um bom atendimento e a agilidade nos processos, que geram bem-estar e atendem as expectativas.

Logo, qualquer banco precisa estar disposto a evitar que seus clientes esperem em filas, transformando tal fato em indicador de bons serviços, e conquistando mercado face à satisfação dos clientes.

Com o intuito de buscar um aperfeiçoamento do desempenho destes sistemas a Teoria das Filas, como uma das técnicas da Pesquisa Operacional (PO), é utilizada, frequentemente, para resolução de problemas que envolvem tempo de espera, ou seja, em um determinado sistema, clientes chegam para serem atendidos, recebem o serviço e depois se retiram do sistema (ROMERO et al. 2010). Ainda, segundo Dávalos (2012) a teoria parte do seguinte pressuposto “a formação de filas excede a capacidade de fornecer determinado serviço” e envolve o estudo matemático utilizando ferramentas de tratamento estatístico ou estocástico.

Frente a este contexto, o objetivo deste trabalho é realizar um estudo que dimensione o sistema de atendimento de uma agência bancária situada no município de Farroupilha - RS, quanto ao seu atendimento, avaliando do ponto de vista operacional e propondo um sistema alterado por meio da criação de cenários para melhoria dos serviços prestados.

2. Pesquisa Operacional

Longaray (2013, p. 4) define a Pesquisa Operacional (PO) como uma “ciência que se dedica exclusivamente ao desenvolvimento de modelos para auxiliar as pessoas e organizações em seus processos decisórios”. A PO surgiu oficialmente após a Segunda Guerra Mundial para resolver problemas de alocação de recursos (ANDRADE, 2015). No período pré-guerra, os militares britânicos concentraram seus esforços no desenvolvimento de armas e equipamentos de apoio, frente a este cenário surgiu a necessidade de sistemas que alocassem tanto recursos de mão-de-obra quanto de materiais (CARTER; PRICE, 2001).

Após o término da guerra, a PO despertou o interesse de aplicação fora do ambiente militar. Na década de 1950 ela passou a ser utilizada nas áreas empresariais, governamentais e comerciais, o que alavancou sua disseminação (HILLIER; LIEBERMAN, 2013). O sucesso de sua aplicação no meio empresarial está relacionado a objetividade de suas técnicas por meio da utilização de modelos que visam traduzir de forma objetiva, clara e estruturada os problemas vivenciados no dia-a-dia de uma organização (LONGARAY, 2013). A utilização de modelos de acordo com Andrade (2015) permite testar uma decisão antes dela ser implementada, o que pode auxiliar na obtenção de uma decisão melhor. O objetivo da PO consiste em encontrar a solução ótima para o modelo que representa o problema estudado (HILLIER; LIEBERMAN, 2013).

2.1. Teoria das Filas

Em 1908, na cidade de Copenhague na Dinamarca, o matemático A. K. Erlang iniciou a abordagem matemática das filas quando trabalhava em uma companhia telefônica estudando problemas de dimensionamento de centrais telefônicas (PRADO, 2006). Arenales *et al.* (2007, p.433) conceitua a Teoria das Filas como “um ramo da Pesquisa Operacional que estuda as relações entre demandas em um sistema e os atrasos sofridos pelos usuários deste sistema”. A Teoria das Filas visa resolver problemas de congestionamento em sistemas, onde “clientes” solicitam “serviços” (ANDRADE, 2015).

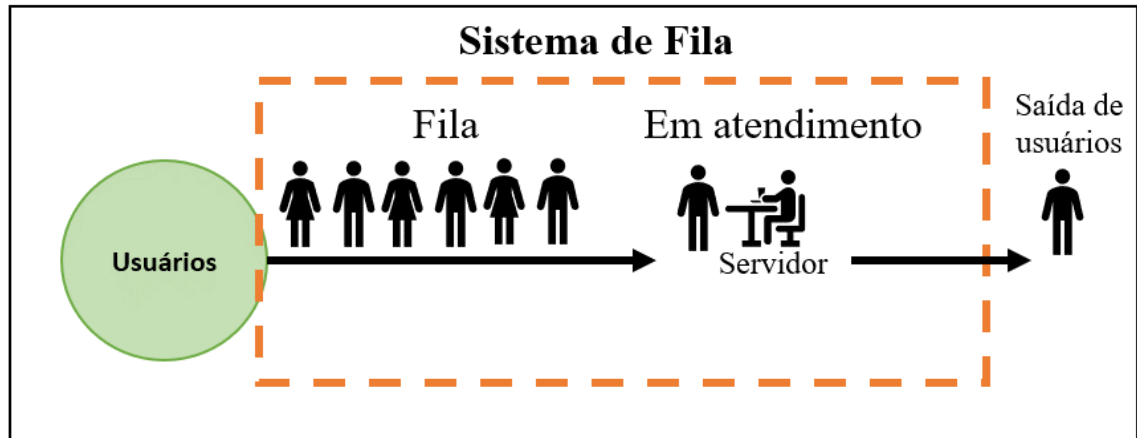
A Teoria das Filas auxilia no balanceamento adequado entre os custos de oferecer determinado serviço e o atraso sofridos pelos clientes do sistema (ARENALES *et al.*, 2007). Segundo Fogliatti e Mattos (2007) a Teoria das Filas estuda sistemas que resultam em esperas, visando determinar e avaliar quantidades, conhecidas como medidas de desempenho, como por exemplo, tamanho da fila, tempo ocioso dos prestadores de serviço, tempo de espera para atendimento.

2.1.1. Sistema de filas

Um sistema de filas é caracterizado por usuários que necessitam de atendimento e que precisam esperar até que o posto de atendimento esteja disponível (ANDRADE, 2015). Importante observar que filas nem sempre são compostas por pessoas, pode ser de carros em um congestionamento, de peças em indústrias e de aviões aguardando para decolar.

De acordo com Fogliatti e Mattos (2007) um sistema de filas é composto por clientes, postos de serviço e por um espaço para espera. A Figura 1 representa uma ilustração de um sistema de filas.

Figura 1 – Ilustração de um sistema de filas



Fonte: os autores (2020)

O sistema de filas pode ser dividido em quatro tipos: fila única e único servidor, como em consultórios médicos por exemplo; fila única com múltiplos servidores em paralelo, em casas lotéricas; múltiplas filas e múltiplos servidores em paralelo, em supermercados e; fila única e múltiplos servidores em série, em restaurantes *fast food*, onde o cliente aguarda para fazer o pedido em uma fila e depois aguarda em outra fila para retirar o pedido (ARENALES *et al.*, 2007).

2.1.2. Disciplina da fila

De acordo com Taha (2008) a disciplina da fila representa a ordem que os clientes serão atendidos em uma fila. Para Fogliatti e Mattos (2007) a disciplina da fila é definida pela gerência do sistema e as mais utilizadas são:

- a) FIFO (*first in – first out*): o primeiro a chegar é o primeiro a sair, ou seja, os clientes são atendidos por ordem de chegada;
- b) LIFO (*last in – last out*): o último cliente a chegar é o primeiro a ser atendido;
- c) PRI (*priority service*): o atendimento segue uma ou mais prioridades definidas;
- d) SIRO (*service in random order*): o atendimento é realizado por uma ordem aleatória.

2.1.3. Processo de chegada

O processo de chegada de usuários em um sistema é determinístico se o número de chegadas e os instantes de tempo das mesmas é conhecido. Caso contrário, o comportamento é aleatório, sendo um processo estocástico definido por uma distribuição de probabilidades (FOGLIATTI; MATTOS, 2007). Para que seja possível caracterizar se o processo de chegadas pode ser considerado uma distribuição de probabilidades o processo de chegadas deve estar em estado estacionário, ou seja, o processo de hoje deve ser o mesmo de amanhã (ANDRADE, 2015).

O processo de chegada da população no sistema pode ser quantificado de duas maneiras: taxa média de chegada e intervalo médio entre as chegadas. A conotação usada para taxa média de chegada (ou ritmo de chegada) é a letra grega “ λ ” e usa-se “IC” para intervalo médio entre chegadas.

2.1.4. Processo de atendimento

O processo de atendimento é caracterizado pelo fluxo de usuários atendidos (FOGLIATTI; MATTOS, 2007). É considerado um processo estacionário e estocástico, onde os tempos de atendimento são orientados por uma distribuição exponencial. Entretanto, o processo depende da disponibilidade dos clientes e das características próprias dos atendentes (ARENALES *et al.*, 2007).

Os valores calculados no processo de atendimento são a duração média de atendimento e o ritmo médio de atendimento, ou seja. Para isto, usa-se a letra grega “ μ ” para determinar o ritmo médio do atendimento e a sigla “TA” para o tempo médio de atendimento (PRADO, 2006).

2.1.5. Notação de Kendal Lee

A notação de Kendall-Lee foi desenvolvida para simplificar a análise dos sistemas de filas. É composta de seis características em que, as três primeiras foram criadas por D. Kendall em 1953 e as três últimas foram adicionadas por A. Lee em 1968 (ARENALES *et al.*, 2007). A notação é A/B/c/K/m/Z e as seis características estão descritas abaixo:

- a) A: distribuição dos intervalos entre as chegadas;
- b) B: distribuição do processo de tempo de serviço;

- c) c: quantidade de atendentes;
- d) K: capacidade máxima do sistema;
- e) m: tamanho da população que fornece clientes para o sistema;
- f) Z: descreve a disciplina da fila.

Também existe uma notação condensada, dada por três termos (A/B/c), em que se supõe que não haja limite para o tamanho da fila, que a população seja infinita e a que disciplina da fila seja FIFO (*first in, first out*), ou seja, o primeiro a entrar é o primeiro a sair (FOGLIATTI; MATTOS, 2007).

2.1.6. Distribuição de Poisson

A distribuição de Poisson é utilizada para descrever o número de acontecimentos que ocorrem em um certo período de tempo ou em uma área ou volume específicos (MCCLAVE; BENSON; SINCICH, 2009). Nesta distribuição os intervalos de tempo entre chegadas e tempo de serviço apresentam uma distribuição exponencial. O modelo baseia-se no estado de equilíbrio, o qual é alcançado quando o sistema opera há um determinado tempo (HILLIER; LIEBERMAN, 2013). A Equação 1 representa a probabilidade de existirem x ocorrências:

$$P(x) = \frac{(\lambda)^x * e^{-\lambda}}{x!} \quad (1)$$

Sendo:

P(x) = probabilidade de x chegadas em t período de tempo;

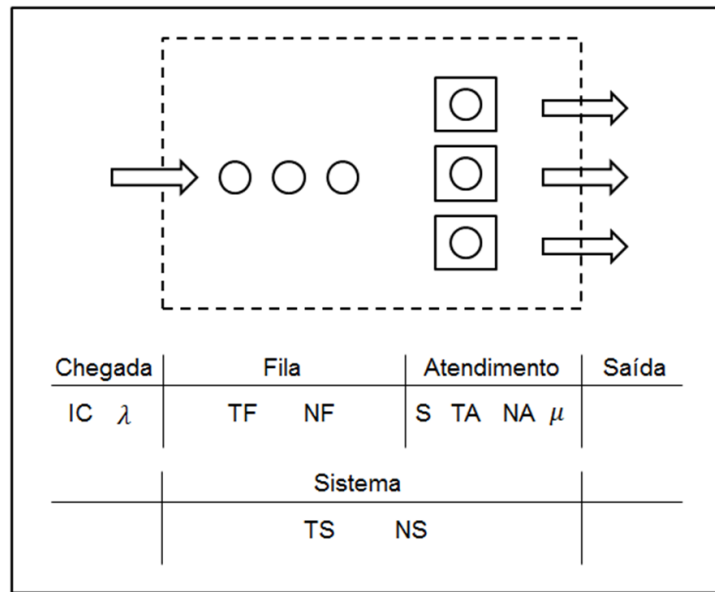
λ = taxa média de chegadas por unidade de tempo;

e = exponencial (2,7183).

2.1.7. Medidas de desempenho do sistema

Por meio da aplicação de Teoria das Filas é possível medir a eficiência de um sistema (FOGLIATTI; MATTOS, 2007). Na Figura 2 é possível observar um sistema de filas com a localização das medidas de desempenho representadas por variáveis aleatórias.

Figura 2 – Sistema de filas



Fonte: Rossa (2014)

As variáveis acima representadas são divididas em variáveis de chegada, da fila, de atendimento e do sistema (PRADO, 2006). No Quadro 1 estão representadas as variáveis e seus significados:

Quadro 1 – Variáveis para as medidas de desempenho

Termo	Descrição
IC	Intervalo entre as chegadas
λ	Taxa média de chegada
TF	Tempo de espera estimado na fila
NF	Número esperado de clientes na fila
C	quantidade de atendentes
TA	Tempo médio de atendimento
NA	Número médio de clientes em atendimento
μ	Taxa média de atendimento
TS	Tempo de espera estimado no sistema
NS	Número esperado de clientes no sistema

Fonte: Os autores (2020)

No Quadro 2 tem-se as equações relativas aos indicadores de Teoria das Filas.

Quadro 2 – Equações relativas aos indicadores de Teoria das Filas

Nome	Equações
Taxa de utilização dos atendentes	$\rho = \frac{\lambda}{\mu}$
Probabilidade do sistema estar ocupado	$\rho = \frac{\lambda}{c \mu}$
Intensidade de tráfego	$i = \frac{ \lambda }{ \mu } = \frac{ TA }{ IC }$
Intervalo entre chegadas	$IC = \frac{1}{\lambda}$
Tempo do atendimento	$TA = \frac{1}{\mu}$
Relações entre fila, sistema e atendimento	$NS = NF + NA$ $NA = \frac{\lambda}{\mu}$ $NS = NF + \frac{\lambda}{\mu} = NF + \frac{TA}{IC}$ $TS = TF + TA$ $NA = \rho = \frac{\lambda}{c \mu}$
Fórmula de Little	$NF = \lambda.TF$ $NS = \lambda.TS$

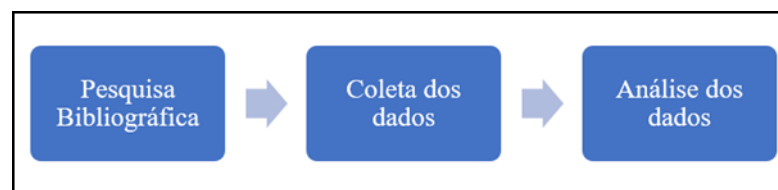
Fonte: Adaptado de Prado (1999) e Andrade (2015)

Todas as equações descritas no Quadro 2 independem da quantidade de atendentes e do modelo de fila, são as equações fundamentais básicas (PRADO, 1999).

3. Metodologia

Este estudo trata-se de um estudo de caso de abordagem quantitativa e de natureza aplicada. Para a realização do mesmo foram realizadas pesquisas em livros e dissertações para obter mais informações sobre o assunto. Na Figura 3 tem-se as etapas deste estudo.

Figura 3 – Etapas do estudo



Fonte: Os autores (2020)

Após a revisão bibliográfica, foi realizada a etapa de coleta dos dados. Para a realização do trabalho foi definido realizar o estudo em uma agência bancária, devido ao fato de este ambiente ser de fácil formação de filas. A coleta dos dados foi realizada no período de 10 dias durante o intervalo de tempo das 11 horas as 16 horas.

Com os dados coletados, realizou-se a análise dos mesmos visando entender se o atual dimensionamento do sistema é satisfatório ou não com o objetivo de propor cenários caso necessário.

3.1. Local escolhido

Este trabalho estudou o processo de chegada dos clientes de uma agência bancária situada na cidade de Farroupilha, Rio Grande do Sul, onde foram realizadas observações visando entender o funcionamento dos caixas de atendimento. Esta agência opera das 11 horas as 16 horas de segunda a sexta-feira sem fechar ao meio dia realizando operações como pagamentos de conta, recebimento de cheques, saques, depósitos entre outros serviços.

Atualmente, o referido ambiente dispõe de três guichês de atendimento, sendo que um deles atende clientes preferenciais quando se faz necessário. Os funcionários trabalham em paralelo e o ambiente conta com uma fila única. Pode-se observar que o sistema obedece à notação Kendall Lee $M/M/3/\infty/FIFO$. Ou seja, os tempos entre chegadas sucessivas assim como os tempos de atendimento seguem uma distribuição exponencial (M/M), há três postos de atendimento (3), o sistema possui capacidade infinita (∞), tendo em vista que não há incidência de impedimento de entrar na agência por superlotação e a disciplina de atendimento é FIFO.

3.2. Dados obtidos

O levantamento dos dados foi realizado diariamente em cada um dos três caixas para atendimento, em arquivo físico e, posteriormente, foram transcritos em arquivo digital no Microsoft Excel. O início da coleta de dados foi no dia 25 de Novembro de 2019 e o término no dia 06 de Dezembro de 2019. Na Tabela 1 tem-se as informações referentes ao tempo médio de atendimento e a quantidade de pessoas que chegaram por hora.

Tabela 1 – Tempo médio de atendimento e número de chegadas

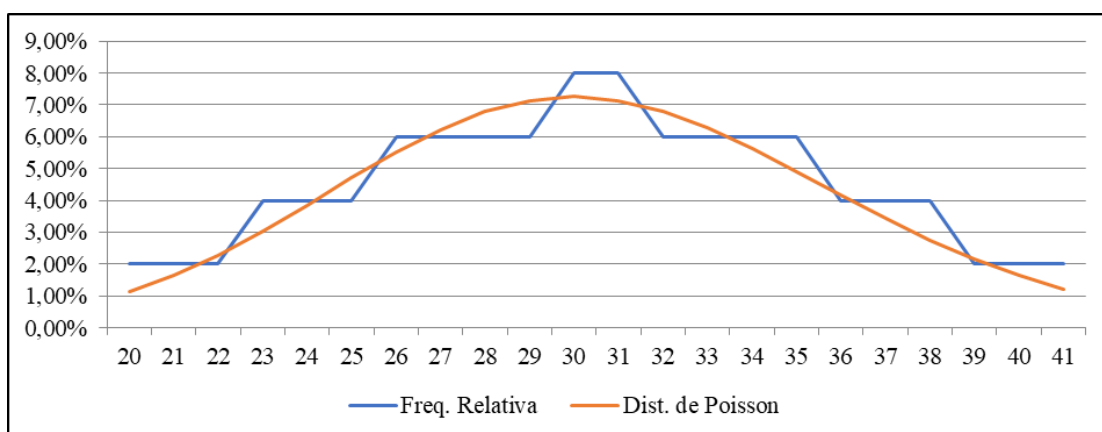
Medida	Valor
--------	-------

Tempo médio de atendimento (minutos)	5,13
Quantidade média de chegadas (pessoas)	30,50

Fonte: Os autores (2020)

No período de 10 dias chegaram 1525 pessoas, gerando uma taxa média de 152,5 clientes por dia e uma taxa média de chegadas (λ) de 30,5 clientes por hora. A curva de frequência relativa observada mostra que a mesma se comporta como uma distribuição de Poisson, conforme Figura 3. No caso, pode-se verificar que o fato de o λ ser relativamente alto, a curva se comporta similar a uma distribuição normal, podendo ser englobada como um caso particular da distribuição de Poisson.

Figura 3 – Frequência relativa X Distribuição de Poisson



Fonte: Os autores (2020)

Para os cálculos estatísticos referentes ao atendimento elaborou-se uma tabela com as durações dos atendimentos realizados pelos três atendentes e em seguida foram listadas informações do menor e do maior valor obtido e a média dos tempos de atendimento. Na Tabela 2 é possível observar estas informações.

Tabela 2 – Dados do atendimento

Dados do atendimento	
Menor tempo (em minutos)	3,10
Maior tempo em minutos	11,98
Média (minutos por cliente)	5,13

Fonte: Os autores (2020)

Nota-se por meio da Tabela 2 que o menor tempo de atendimento foi de 3,10 minutos e o maior tempo foi de 11,98 minutos, com desvio padrão de 1,28 minutos. O maior valor de atendimento observado justifica-se pelo fato de que um mesmo cliente pode necessitar de serviços de pagamento de contas, saques e depósitos de dinheiro, por exemplo.

4. Resultados

Por meio da aplicação das equações correspondentes e os dados e indicadores da Teoria das Filas, presentes no Quadro 2, referentes ao dimensionamento do sistema empregado no ambulatório estudado, foi possível chegar aos resultados apresentados na Tabela 3.

Tabela 3 – Parâmetros do sistema

Parâmetro	Resultado	Unidade
Taxa média de chegada (λ)	30,50	pessoas/ hora
Número médio de clientes que aguardam na fila (NF)	5,83	pessoas
Tempo médio de permanência na fila (TF)	0,19	hora
Intervalo médio entre chegadas (IC)	0,03	horas/ pessoa
Tempo médio de atendimento (TA)	0,03	horas
Número médio de clientes que estão sendo atendidos (NA)	0,87	pessoas
Taxa de utilização (ρ)	87,01	% (por cento)
Tempo médio de permanência no sistema (TS)	0,22	horas
Número médio de clientes no sistema (NS)	6,70	pessoas

Fonte: Os autores (2020)

A taxa de ocupação do sistema é de 87% com três atendentes, o que explica a fila observada em alguns momentos durante o período da análise. Os dados apresentados na Tabela 3 mostram que o sistema é estável, pois a taxa média de atendimento geral é maior que a taxa de chegada. A Tabela 4 representa uma simulação considerando que o sistema utilizasse quatro atendentes.

Tabela 4 – Parâmetros do sistema considerando quatro atendentes

Parâmetro	Resultado	Unidade
Taxa média de chegada (λ)	1,23	pessoas/ hora
Número médio de clientes que aguardam na fila (NF)	0,04	pessoas

Tempo médio de permanência na fila (TF)	0,03	hora
Intervalo médio entre chegadas (IC)	0,02	horas/ pessoa
Tempo médio de atendimento (TA)	0,65	horas
Número médio de clientes que estão sendo atendidos (NA)	65,26	pessoas
Taxa de utilização (ρ)	0,06	% (por cento)
Tempo médio de permanência no sistema (TS)	1,88	horas
Número médio de clientes no sistema (NS)	1,23	pessoas

Fonte: Os autores (2020)

Considerando um atendente a mais do que o sistema tem atualmente, a taxa de ocupação seria de 65%, o que tornaria o sistema mais ágil para o cliente nos dias e horários de maiores picos no sistema, porém geraria mais ociosidade para os atendentes além de custos maiores por ter um atendente a mais. Na configuração com quatro atendentes a agência bancária teria uma redução do percentual de ocupação no valor de 25%.

5. Conclusões

O estudo de Teoria das Filas mesmo sendo breve para o caso analisado apresenta dados importantes que auxiliam na tomada de decisão da agência bancária. O dimensionamento atual do banco é satisfatório para atender a demanda, uma vez que a taxa de ocupação é de 87%, o que explica a fila observada em alguns períodos da análise. Entretanto, quando comparado com um cenário considerando um atendente a mais visando agilizar o atendimento aos clientes, evitando a formação de filas, a taxa de ocupação é de 65%.

Com quatro atendentes o sistema torna-se mais ágil, porém cabe a gestão do banco avaliar em termos de custo, pois o sistema terá que ter um funcionário a mais. Todavia, por mais que o sistema seja estável é necessário uma avaliação periódica do mesmo para garantir a satisfação dos clientes com os serviços prestados.

REFERÊNCIAS

ALBRECHT, Karl; BRADFORD, Lawrence J. **Serviços com Qualidade: a vantagem Competitiva**. São Paulo: Makron Books, 1992.

ANDRADE, Eduardo Leopoldino. **Introdução à pesquisa operacional: métodos e modelos para análise de decisões**. 5. ed. Rio de Janeiro: LTC, 2015.

ARENALES, Marcos; ARMENTANO, Vinícius; MORABITO, Reinaldo; YANASSE, Horacio. **Pesquisa operacional**. Rio de Janeiro: Elsevier, 2007.

CARTER, Michael W.; PRICE, Camille C.. **Operations research: a practical introduction**. Boca Raton: CRC Press, 2001.

DÁVALOS, Pablo Bezzera. **Hidroquímica do Estuário do Rio Caravelas**, Caravelas – BA / Pablo Bezzera Dávalos – Natal, RN, 2012.

FOGLIATTI, M. C.; MATTOS, N. M. C. **Teoria de filas**. 8. Ed. Rio de Janeiro: Editora Interciência, 2007.

GIANESI, Irineu G. N.; CORRÊA, Henrique Luis. **Administração estratégica de serviços**. São Paulo. Atlas, 1996.

HILLIER, Frederick S.; LIEBERMAN, Gerald J.. **Introdução à pesquisa operacional**. 9. ed. Porto Alegre: AMGH, 2013.

LONGARAY, André Andrade. **Introdução à pesquisa operacional**. São Paulo: Saraiva, 2013.

MCCLAVE, James T., BENSON, P G., SINCICH, Terry. **Estatística para administração e economia**. São Paulo: Pearson, 2009.

PRADO, Darci. **Teoria das filas e da simulação**. 3ª ed. Belo Horizonte: INDG TecS, 2006.

RICARDO, Fabiana Alves. **Gestão da qualidade total: a qualidade como valor percebido pelo cliente**. 2010. 63 f. TCC (Graduação) - Curso de Administração Com Habilitação em Comércio Exterior, Faculdade Instituto de Ensino Superior de Bauru, Bauru, 2010.

ROMERO, C. M.; SALES, D. S.; VILAÇA, L. L.; CHAVEZ, J. R. A.; CORTES, J. M. (2010). **Aplicação da teoria das filas na maximização do fluxo de paletes em uma indústria química**. Pesquisa Operacional para o Desenvolvimento, v.2, n.3, p.226-231.

ROSSA, Deise. **Proposta de metodologia de aplicação da teoria das filas em um sistema ambulatorial**. 2014. 91 f. Dissertação (Graduação em Engenharia de Produção) – Universidade de Caxias do Sul. Caxias do Sul, 2014.

TAHA, Hamdy A.. **Pesquisa operacional: uma visão geral**. 8. ed. São Paulo: Pearson Prentice Hall, 2008.