

Dimensionamento de Recursos em Maternidades Utilizando Modelo de Perda de Erlang

Ricardo Bruno Osés de Oliveira
Flávio Geraldo Coelho Rocha

Escola de Engenharia Elétrica, Mecânica e de Computação - UFG
Av. Universitária, n.º 1488 - quadra 86 - Setor Leste Universitário – Goiânia – Goiás
ricfire.ricardo@gmail.com, flaviogcr@ufg.br

RESUMO

Neste artigo foi considerado a aplicação do modelo de perda de Erlang $M/G/m/m$ para descrever o fluxo de gestantes que chegam às maternidades. Utilizou-se os dados de duas maternidades que descrevem o processo de chegadas e informam o tempo de permanência e a capacidade de internação em leitos para partos. O modelo de Erlang foi então utilizado para o dimensionamento da capacidade das duas maternidades analisadas. Foi possível estimar o número requerido de leitos variando-se as taxas de admissões de gestantes recusadas, a demanda exigida e o tempo de permanência das gestantes em cada uma das duas maternidades.

PALAVRAS CHAVE. Maternidade, Dimensionamento, Erlang-B.

PO na Área de Saúde

ABSTRACT

In this paper, the application of the Erlang $M/G/m/m$ loss model was considered to describe the flow of pregnant women arriving at maternity hospitals. Data was used from two maternity hospitals describing the process of arrivals and reporting length of stay and inpatient bed capacity for deliveries. The Erlang model was then used to size the capacity of the two maternity hospitals analyzed. It was possible to estimate the required number of beds by varying the admission rates of refused pregnant women, the required demand, and the length of stay of pregnant women in each of the two maternity hospitals.

KEYWORDS. Maternity. Dimensioning. Erlang-b.

OR in Healthcare

1. Introdução

O aumento do número de nascimentos, a diminuição dos leitos obstétricos e a consequente superlotação nas unidades perinatais levanta questões relativas ao planejamento, previsão e o cálculo do tamanho adequado para estas unidades. A indisponibilidade de leitos obstétricos, de leitos de UTI e de uma estrutura mínima que ofereça assistência às parturientes pode levar à rejeição de internação dessas pacientes e a necessidade de transferências para outras maternidades. Uma gestante rejeitada de uma unidade perinatal pode transbordar para outra unidade na mesma rede. Por conseguinte, uma maternidade pode receber tanto as suas próprias pacientes como as pacientes que transbordam de outras maternidades.

Como a demanda por serviços nas unidades perinatais aumenta significativamente, tem-se como alternativa aumentar o número de leitos ou adequar a demanda ao número de leitos disponíveis, o que muitas vezes só é obtido à custa de baixas taxas de ocupação [Jones, 2012]. Diante da diminuição dos leitos obstétricos a melhor opção seria adequar o número de partos e atendimentos à capacidade física da unidade perinatal. Geralmente, algumas maternidades elaboram, ao início de cada ano, um relatório de produção estimada para cada mês buscando regular o número de partos que poderão ser realizados, o número de procedimentos cirúrgicos e o número de atendimentos e demais serviços oferecidos. No entanto, as taxas de superlotação estão aumentando devido ao mal planejamento da disponibilização dos recursos existentes.

As decisões sobre o planejamento da capacidade nos hospitais são tomadas geralmente sem o aporte de análises quantitativas baseadas em estudos e modelos [Laffel e Blumenthal, 1989]. O que explica a diminuição do número de leitos obstétricos e o aumento nas taxas de ocupação das unidades perinatais. A variabilidade no tempo de permanência das gestantes nas unidades perinatais causam um grande impacto no funcionamento diário dessas unidades e nos requisitos de capacidade. Se esta variabilidade for desconsiderada durante a modelagem, surgirá uma representação irrealista e estática da realidade [De Bruin et al., 2007].

O planejamento e o dimensionamento dos recursos em unidades hospitalares são amplamente abordados em diversos trabalhos sob a perspectiva da teoria de filas. [Green et al., 2006], [De Bruin et al., 2007], [Tseytlin, 2009] [de Bruin et al., 2010], [Yankovic e Green, 2011], [Véricourt e Jennings, 2011], [Pehlivan et al., 2012], [Mandelbaum et al., 2012], [Pehlivan, 2014], [Armony et al., 2015]. Diversos autores já estudaram o modelo de filas de espera com o uso da equação de *Erlang-B* para o dimensionamento de recursos em unidades hospitalares [Green e Nguyen, 2001], [Harper e Shahani, 2002], [Gorunescu et al., 2002], [Kokangul, 2008], [Li et al., 2009], [de Bruin et al., 2010], [Belciug e Gorunescu, 2015], [Essen et al., 2015]. A equação de *Erlang-B* possui variáveis importantes que possibilitam o cálculo de várias medidas de interesse para o planejamento tanto de recursos humanos como da capacidade física dos serviços hospitalares.

O restante do artigo está organizado da seguinte forma: a seção 2 descreve o modelo matemático utilizado; a seção 3 descreve o modelo estrutural do fluxo de gestantes nas maternidades; por fim, na seção 4 são apresentados os resultados obtidos e as considerações finais são descritas na seção 5.

2. Modelo Matemático

As pacientes de redes perinatais são urgentes, e, portanto, uma capacidade de atendimento inadequada em uma maternidade faz com que a gestante espere muito ou seja transferida para outra unidade, o que pode causar longas distâncias de deslocamento para outra unidade de saúde resultando em rejeição e transbordamento de gestantes. Esse problema pode e deve ser evitado por

meio de um apropriado dimensionamento e operação eficiente dos recursos disponíveis Pehlivan et al. [2013].

Baseado na teoria de filas muitos estudos analisam modelos de atraso que são aplicados para problemas de planejamento de capacidade em unidades hospitalares [Green e Nguyen, 2001], [Green, 2002],[de Bruin et al., 2010]. O fluxo dos pacientes que entram em uma fila de espera é analisado para quantificar a probabilidade de rejeição e a probabilidade de atraso, que são importantes medidas de desempenho para um sistema de enfileiramento. O objetivo principal é encontrar soluções para um melhor dimensionamento e adequação dos recursos existentes na unidade hospitalar. Como em uma unidade perinatal a rejeição de gestantes é um problema real a ser evitado, um modelo de enfileiramento adequado é o $M/G/m/m$ que incorpora o bloqueio aos clientes que entram no sistema e encontram todos os servidores ocupados.

2.1. Modelo de Enfileiramento $M/G/m/m$

Considere o sistema de enfileiramento $M/M/m/m$ descrito na Figura 1. Os clientes chegam de acordo com um processo de Poisson. O tempo entre as chegadas obedece a uma distribuição exponencial e os tempos de serviço são exponencialmente distribuídos. Uma fila $M/M/m/m$ trata-se então de um sistema com m servidores e capacidade m , sem fila de espera. Portanto, se os m servidores estiverem todos ocupados quando um novo cliente chegar, ele será perdido pois o sistema está cheio, ocorrendo então o bloqueio do sistema.

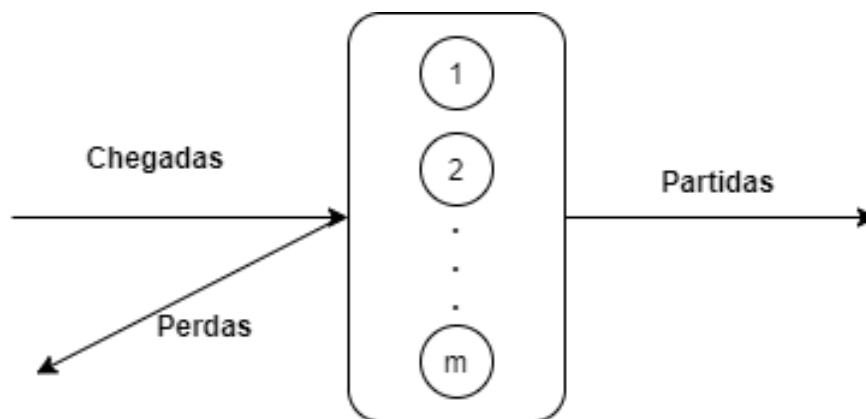


Figura 1: Sistema de enfileiramento $M/M/m/m$.

O estado do sistema é definido pelo número de clientes presentes no sistema e o espaço de estado é finito. O estado pode ser descrito por um processo de nascimento e morte para qual o diagrama de transição de estados é mostrado na Figura 2.

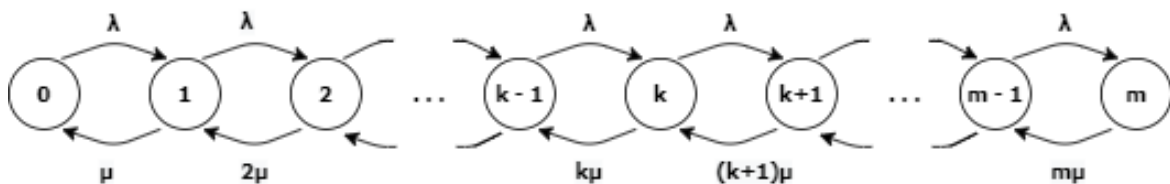


Figura 2: Diagrama de transição estados para um sistema $M/M/m/m$.

Seja p_k a probabilidade de haver k clientes no sistema em um momento arbitrário em equilíbrio, onde $k = 0, 1, 2, \dots, m$. As probabilidades de estado em equilíbrio para $\{p_k; 0 \leq k \leq m\}$ podem ser expressas como:

$$\begin{aligned} \mu p_1 &= \lambda p_0, \\ \lambda p_{k-1} + (k+1)\mu p_{k+1} &= (\lambda + k\mu)p_k, \quad 1 \leq k \leq m, \\ \lambda p_{m-1} &= m\mu p_m. \end{aligned} \quad (1)$$

Isto é equivalente ao conjunto:

$$k\mu p_k = \lambda p_{k-1}, \quad 1 \leq k \leq m, \quad (2)$$

onde:

$$p_k = \frac{\lambda}{k\mu} p_{k-1} = \frac{\lambda^2}{k(k-1)\mu^2} p_{k-2} = \dots = \frac{1}{k!} \left(\frac{\lambda}{\mu}\right)^k p_0, \quad 1 \leq k \leq m. \quad (3)$$

A intensidade do tráfego ou o fator de utilização ρ é definido como:

$$\rho = \frac{\lambda}{\mu}. \quad (4)$$

Encontrando p_0 através da condição de normalização:

$$\sum_{k=0}^m p_k = 1, \quad (5)$$

obtem-se a distribuição de probabilidade para o número de clientes no sistema:

$$p_k = p_0 \frac{\rho^k}{k!} = \frac{\rho^m / m!}{\sum_{k=0}^m \rho^k / k!}, \quad 1 \leq k \leq m. \quad (6)$$

Esta é a distribuição de Poisson truncada. Os clientes que chegam quando todos os servidores estão ocupados são bloqueados. A probabilidade de haver k clientes no sistema imediatamente antes de uma chegada é igual à probabilidade de haver k clientes no sistema em um tempo arbitrário dado na Equação (6) [Wolff, 1989]. Portanto, a probabilidade de bloqueio, também conhecida como equação de *Erlang-B* é dada por:

$$P_m = \frac{(\lambda/\mu)^m / m!}{\sum_{k=0}^m (\lambda/\mu)^k / k!} \quad (7)$$

Pode-se provar que também para uma distribuição geral do tempo de serviço as probabilidades P_n são dadas pela Equação (7), onde $\rho = \lambda\mu$ é frequentemente referido como a carga oferecida ao sistema. Obtém-se então a equação de *Erlang-loss* para o modelo M/G/m/m:

$$P_m = \frac{(\lambda\mu)^m / m!}{\sum_{k=0}^m (\lambda\mu)^k / k!} \quad (8)$$

3. Descrição do Modelo

Considere o modelo estrutural do fluxo de gestantes em uma maternidade mostrado na Figura 3. As gestantes que chegam ao sistema são admitidas se houver leitos disponíveis. Caso a gestante encontre todos os leitos ocupados ela é recusada e deixa o sistema. Na prática, uma admissão recusada pode resultar em um desvio para outra maternidade não-preferível.

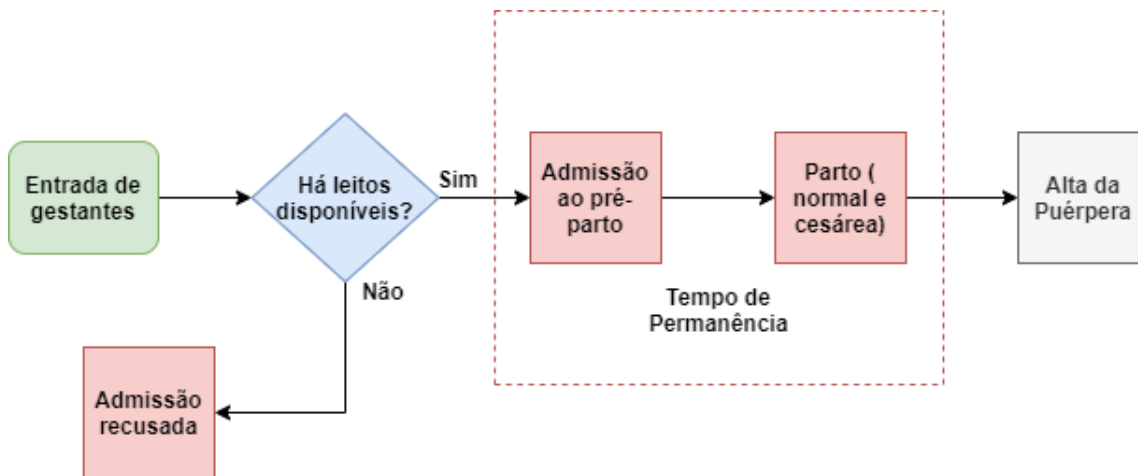


Figura 3: Modelo estrutural do fluxo de gestantes em uma maternidade.

O número de admissões recusadas pode ser interpretado como um indicador do nível de serviço e é importante para a qualidade do atendimento [de Bruin et al., 2010]. As maternidades mantêm registro do número de admissões contabilizada como o número de procedimentos realizados, aqui considerado, como o número de partos. O tempo de permanência é contabilizado desde a entrada da gestante na maternidade até o momento da alta, mesmo que essa seja transferida para outra ala da maternidade. O tempo de permanência é uma métrica essencial para evitar congestionamento e atrasos nos atendimentos.

A capacidade de uma maternidade é medida em termos de leitos operacionais e a quantidade de pessoal disponível é baseada nessa capacidade operacional. Isto é feito através de uma relação de pessoal por leito operacional. A rejeição de gestantes também pode acontecer em outros ambientes da maternidade após a admissão e o parto. Há casos em que as puérperas necessitam de cuidados intensivos devido a complicações durante e ou antes do parto. Isso é um agravante pois, com as gestantes em estado crítico de saúde a transferência para outra unidade que possua leitos para tratamento intensivo se torna algo delicado e não preferível.

As unidades perinatais exigem um certo número de leitos para admitir e tratar todas as gestantes. O número necessário de leitos depende do tempo de permanência de cada gestante e do número de internações e atendimentos por dia. A taxa de admissões e de perdas são variáveis aleatórias. Caso fossem variáveis determinísticas, ou seja, a cada dia exatamente λ gestantes chegassem com um tempo médio de permanência $1/\mu$, então a cada dia haveria exatamente $\lambda/\mu = \rho$ leitos ocupados quando não ocorresse nenhum bloqueio. Entretanto, na prática, o número de gestantes e a média de permanência são variáveis estocásticas, já que o tempo de permanência é diferente para cada gestante e o número de internações varia por dia. Para garantir que haja leitos suficientes em um ambiente estocástico, certamente precisa-se de mais leitos do que ρ para lidar com a flutuação do número de gestantes admitidas. Porém, na prática, nem sempre será possível admitir todas as gestantes, porque então o número de leitos necessários teria que ser extremamente grande e uma

grande parte dos leitos ficariam ociosos na maior parte do tempo perdendo eficiência na utilização dos recursos [Essen et al., 2015].

Com os dados obtidos das duas maternidades anônimas de Goiânia pode-se usar a Equação (8) de perda de *Erlang* para determinar a média de permanência e a taxa de rejeição de gestantes, o número de leitos necessários para se atingir uma taxa de rejeição predefinida e a taxa média de admissões desejada para a quantidade de leitos disponíveis. Na Equação (8) tem-se que:

- λ = número médio chegadas por dia;
- μ = tempo médio de permanência das gestantes;
- n = número de leitos disponíveis;
- P_n = taxa de rejeição de gestantes.

Usando a fórmula de Little [Kleinrock, 1975] a taxa de ocupação é obtida fazendo:

$$T_o = \frac{\lambda\mu(1 - P_n)}{n}. \quad (9)$$

O número médio de leitos ocupados pode ser encontrado por:

$$\bar{N} = \lambda\mu(1 - P_n). \quad (10)$$

Substituindo P_n em (10) usando (8), obtém-se a expressão para o número médio de leitos ocupados:

$$\bar{N} = \lambda\mu \left(1 - \frac{(\lambda\mu)^n/n!}{\sum_{k=0}^n (\lambda\mu)^k/k!}\right). \quad (11)$$

O número médio de leitos ocupados pode ser obtido nos dados informados pelas duas maternidades e o número de leitos m e o tempo médio de permanência são variáveis conhecidas.

4. Resultados e Discussões

Os registros do número de partos ocorridos nas duas maternidades apresentados na Tabela 1 foram usados para quantificar o número de admissões e a taxa de ocupação. A taxa de ocupação permite avaliar o grau de utilização dos leitos operacionais na maternidade, medir o perfil de utilização e gestão desses leitos e está diretamente relacionado ao intervalo de substituição e à média de permanência. A taxa de ocupação recomendável situa-se entre 75 e 85%. A taxa de ocupação abaixo de 75% indica baixa utilização e ineficiência na gestão hospitalar. Para as duas maternidades analisadas, a média das taxas de ocupação situou-se entre 80 e 90%.

Tabela 1: Número médio estimado de procedimentos por dia.

Tipo de serviço	Maternidade 1	Maternidade 2
Atendimentos	31,4	14,2
Partos normais	7,9	5,4
Partos cirúrgicos (cesáreas)	5,3	3,2

A Tabela 2 apresenta o número de leitos operacionais nas duas maternidades, as médias de permanência observadas e o número de chegadas diárias estimadas no período analisado. A medida

que os leitos vão sendo ocupados, o número de leitos operacionais disponíveis é obtido através da taxa de ocupação. O tempo médio de permanência representa o período médio em dias que as gestantes permanecem internadas na maternidade. Muitos fatores estão associados ao intervalo de tempo de internação, como o tipo do parto e a classificação de risco. No SUS observou-se, para o período de 2008 a 2012, uma média de permanência para partos normais de 2,0 dias e para partos cesáreos de 2,6 dias [da Saúde, 2012].

Tabela 2: Número de leitos, média de permanência e número de chegadas diárias.

Descrição	Maternidade 1	Maternidade 2
Leitos obstétricos (parto normal e cesárea)	47	30
Média de Permanência (dias)	3,5	3
Números de chegadas diárias (λ)	13,86	9,03

Com esses dados, o modelo de perda de Erlang pode então ser usado para determinar o número de leitos necessários quando a média de permanência, a taxa de chegadas e a probabilidade de bloqueio são dadas. Assim, um valor aceitável para a probabilidade de bloqueio pode ser escolhida. Considerando os dados apresentados na Tabela 1 estima-se a taxa de chegadas nas duas maternidades. Como o número de admissões recusadas não é registrada, considerou-se que 5% das gestantes que chegam à maternidade são rejeitadas. Portanto, a demanda diária estimada de gestantes que chegam às duas maternidades, entre admissões recusadas ou não, foi de 13,86 chegadas por dia para a primeira maternidade e 9,03 chegadas por dia para a segunda maternidade. Vale ressaltar que, a demanda considerada foi apenas para os leitos para partos (normal ou cesáreas).

Foram feitas simulações de diversos cenários, reduzindo a demanda nessas duas maternidades para diferentes taxas de rejeição. O objetivo é encontrar o menor número de leitos para os quais a taxa de bloqueio é menor ou igual a taxa fixada. Para isso foram utilizados os tempos médios de permanência reais informado pelas maternidades e os tempos calculados pelos algoritmos de balanceamento de carga em [Oliveira e Rocha, 2020] apresentados na Tabela 4.

Tabela 3: Tempo médio de permanência simulado

Algoritmo	Maternidade 1	Maternidade 2	Média do Sistema
L1	1,08	1,32	1,20
L2	1,15	1,32	1,23
L5	1,47	1,61	1,54
L10	2,19	1,75	1,97
PR	1,65	2,36	2,0
R	1,16	1,53	1,34
T	1,29	1,93	1,56
TP	1,13	1,40	1,26

A Tabela 4 apresenta o número de leitos requeridos para diferentes taxas de rejeição obtidos de acordo com os tempos de permanência apresentados na Tabela 4.

Tabela 4: Número de leitos necessários para diferentes taxas de rejeição na Maternidade M₁.

Número de chegadas diárias (λ)	Tempo médio de permanência (dias)	Nº leitos requeridos para rejeição de:					
		25%	20%	15%	10%	5%	2%
13,86*	3,5*	39	42	45	49	54	59
	1,65 (PR)	19	21	23	25	28	31
	1,16 (R)	14	15	17	18	21	23
	1,29 (T)	15	17	18	20	23	25
	1,13 (TP)	14	15	16	18	20	22
	1,47 (L5)	17	19	20	22	25	28
	2,19 (L10)	25	27	29	32	35	39
	1,15 (JSQ)	14	15	16	18	20	23
	1,10 (HBB-LB)	13	14	16	17	20	22
11,1(-20%)	3,5*	31	34	36	40	44	48
	1,65 (PR)	16	17	19	20	23	25
	1,16 (R)	11	12	13	15	17	19
	1,29 (T)	12	13	15	16	19	21
	1,13 (TP)	11	12	13	14	16	19
	1,47 (L5)	14	15	17	19	21	23
	2,19 (L10)	20	22	24	26	29	32
	1,15 (JSQ)	11	12	13	15	17	19
	1,10 (HBB-LB)	11	12	13	14	16	18

*Valores reais estimados

O tempo médio permanência das gestantes na maternidade M₁ é considerada alta em comparação ao tempo médio de permanência informado pelo SUS, mesmo considerando que o tempo médio de permanência para gestações de alto risco elevem a taxa de permanência global da maternidade. O tempo médio de permanência nacional considerando ambos os partos, em gestações de baixo e alto risco, foi de 2,3 dias em um período de 4 anos [da Saúde, 2012]. Dado esse elevado tempo de permanência, para se obter taxas de admissões recusadas menores é necessário aumentar o número de leitos ou que a demanda de gestantes diminua cerca de 20%. No entanto, dificilmente a demanda teria uma diminuição tão significativa.

Para a demanda e a média de permanência reais informadas pela maternidade, a taxa de rejeição é de 12%, o que significa que aproximadamente a cada 9 gestantes que procuram o serviço para parto, uma não encontra leito disponível para internação. Qual a taxa de admissões recusadas é razoável e, portanto, aceitável ainda é alvo de grandes discussões, no entanto, alguns especialistas em políticas hospitalares muitas vezes buscam atingir uma meta de 5% [de Bruin et al., 2010]. A Tabela 4 revela que, para atingir uma meta de 5% de admissões de gestantes recusadas seriam necessários 54 leitos, ou seja, 7 leitos a mais, dado que a maternidade possui 47 leitos operacionais disponíveis. Caso a demanda diminua em 20% seria necessário apenas mais um leito para que a meta de 5% seja alcançada.

Analisando os resultados obtidos para os tempos médios de permanência calculados pelos algoritmos de balanceamento de carga, observa-se que, à medida que a média de permanência das gestantes cai o número necessário de leitos operacionais cai significativamente. Os tempos médios de permanência obtidos pelos algoritmos situam-se entre 1,10 a 2,19 dias. Hoje no Brasil a exigência mínima do Ministério da Saúde para permanência da parturiente na maternidade é de 24

horas, isso se, o parto tenha sido normal e não teve complicações antes, durante ou após o parto. Portanto, se o tempo médio se mantivesse em torno do tempo mínimo exigido pelo Ministério da Saúde, o planejamento e a gerência dos recursos nas unidades perinatais seria mais facilmente sustentado, pois poderia haver alocação e transferência de leitos entre unidades perinatais próximas, evitando a superlotação de uma maternidade pequena e a ociosidade em uma maternidade de grande porte.

As simulações também foram feitas para a maternidade M_2 . A Tabela 5 apresenta o número de leitos requeridos para diferentes taxas de rejeição. A maternidade M_2 possui uma quantidade menor de leitos em relação a M_1 , além de apresentar uma demanda menor por serviços de parto. Com o tempo médio de permanência de 3 dias como informado, é possível alcançar a meta de 5% de admissões recusadas aumentando apenas 2 leitos. Esse resultado revela que, a maternidade consegue manter uma taxa de admissões recusadas baixa com os recursos disponíveis sem a necessidade de aumento significativo no número de leitos ou diminuição da demanda. Com uma demanda 20% menor, os leitos disponíveis são mais que suficientes, diminuindo a taxa de admissões recusadas a 2%. Assim como na maternidade M_1 , à medida que o tempo de permanência das gestantes diminui, também diminui a necessidade de um número maior de leitos para partos.

Tabela 5: Número de leitos necessários para diferentes taxas de rejeição na Maternidade M_2 .

Número de chegadas diárias (λ)	Tempo médio de permanência (dias)	Nº leitos requeridos para rejeição de:					
		25%	20%	15%	10%	5%	2%
9,03*	3*	22	24	26	29	32	35
	2,36 (PR)	18	19	21	23	26	29
	1,53 (R)	12	13	14	16	18	20
	1,93 (T)	15	16	18	19	22	24
	1,40 (TP)	11	12	13	15	17	19
	1,61 (L5)	13	14	15	17	19	21
	1,75 (L10)	14	15	16	18	20	23
	1,88 (JSQ)	14	16	17	19	21	24
	1,34 (HBB-LB)	11	12	13	14	16	18
7,2(-20%)	3*	18	20	21	23	26	29
	2,36 (PR)	14	16	17	19	21	24
	1,53 (R)	10	11	12	13	15	17
	1,93 (T)	12	13	14	16	18	20
	1,40 (TP)	9	10	11	12	14	16
	1,61 (L5)	10	11	12	14	16	18
	1,75 (L10)	11	12	13	15	17	19
	1,88 (JSQ)	12	13	14	16	18	20
	1,34 (HBB-LB)	9	10	11	12	14	15

*Valores reais estimados

Como nem sempre é possível aumentar de imediato a capacidade das maternidades, a falta de leitos trará restrições às diversas etapas de realização dos partos levando a uma sobrecarga de gestantes aguardando por leitos, sejam anteriores ou até mesmo após o parto. Uma boa alternativa para suprir a dificuldade em se alterar o número de leitos é adequar a quantidade de partos à capacidade física das maternidades.

5. Conclusão

Neste trabalho, foi considerado a aplicação do modelo de perda de Erlang $M/G/m/m$ para descrever o fluxo de gestantes que chegam às maternidades e estimar o número requerido de leitos variando-se as taxas de admissões de gestantes recusadas, a demanda exigida e o tempo de permanência das gestantes em cada uma das duas maternidades. Para estimar o número requerido de leitos foram utilizados os tempos de permanência apresentados em [Oliveira e Rocha, 2020].

Os resultados obtidos pelo modelo trouxeram uma descrição precisa do número de leitos ocupados e do número de leitos necessários de acordo com uma demanda apresentada. Através dos resultados foi possível notar uma importante relação entre a capacidade de uma unidade perinatal, o tempo médio de permanência das gestantes e a probabilidade de uma admissão ser recusada. Em unidades onde o número de leitos é pequeno em comparação a uma elevada demanda por serviços perinatais a probabilidade de admissões recusadas tende a ser alta, comprometendo o gerenciamento de tais unidades.

Com os tempos de permanência obtidos pelos algoritmos de balanceamento de carga em [Oliveira e Rocha, 2020], foi possível alcançar uma maior eficiência na utilização dos leitos operacionais, garantindo baixas taxas de rejeição e uma capacidade operacional menor, o que contribui para diminuição de custos. Por outro lado, unidades com uma capacidade operacional grande e uma demanda baixa leva a ociosidade de recursos ocasionando maiores custos e uma baixa eficiência da unidade. No entanto, o modelo permitiu equilibrar o número de leitos operacionais de acordo com demanda exigida facilitando assim uma boa opção de provisão de leitos, e alto acesso ao serviço. Assim, é possível calcular o número ideal de leitos, dado um baixo nível de rejeição de gestantes. Como trabalho futuro, pretende-se estender o modelo a um conjunto maior de maternidades.

Referências

- Armony, M., Israelit, S., Mandelbaum, A., Marmor, Y. N., Tseytlin, Y., e Yom-Tov, G. B. (2015). On patient flow in hospitals: A data-based queueing-science perspective. *Stochastic systems*, 5 (1):146–194.
- Belciug, S. e Gorunescu, F. (2015). Improving hospital bed occupancy and resource utilization through queueing modeling and evolutionary computation. *Journal of biomedical informatics*, 53: 261–269.
- da Saúde, M. (2012). Ministério da saúde. tabnet win32 3.0: Internações hospitalares do sus - por local de internação - sus. <http://tabnet.datasus.gov.br/cgi/tabcgi.exe?sih/cnv/sxuf.def>. Acessado: 14-02-2020.
- de Bruin, A. M., Bekker, R., Van Zanten, L., e Koole, G. (2010). Dimensioning hospital wards using the erlang loss model. *Annals of Operations Research*, 178(1):23–43.
- De Bruin, A. M., Van Rossum, A., Visser, M., e Koole, G. (2007). Modeling the emergency cardiac in-patient flow: an application of queueing theory. *Health Care Management Science*, 10(2): 125–137.
- Essen, J. T., Houdenhoven, M., e Hurink, J. L. (2015). Clustering clinical departments for wards to achieve a prespecified blocking probability. *OR spectrum*, 37(1):243–271.

- Gorunescu, F., McClean, S. I., e Millard, P. H. (2002). A queueing model for bed-occupancy management and planning of hospitals. *Journal of the Operational Research Society*, 53(1): 19–24.
- Green, L. V. (2002). How many hospital beds? *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, 39(4):400–412.
- Green, L. V. e Nguyen, V. (2001). Strategies for cutting hospital beds: the impact on patient service. *Health services research*, 36(2):421.
- Green, L. V., Soares, J., Giglio, J. F., e Green, R. A. (2006). Using queueing theory to increase the effectiveness of emergency department provider staffing. *Academic Emergency Medicine*, 13(1): 61–68.
- Harper, P. R. e Shahani, A. (2002). Modelling for the planning and management of bed capacities in hospitals. *Journal of the Operational research Society*, 53(1):11–18.
- Jones, R. (2012). A simple guide to a complex problem—maternity bed occupancy. *British Journal of Midwifery*, 20(5):351–357.
- Kleinrock, L. (1975). *Queueing systems. Volume I: theory*. wiley New York.
- Kokangul, A. (2008). A combination of deterministic and stochastic approaches to optimize bed capacity in a hospital unit. *Computer methods and programs in biomedicine*, 90(1):56–65.
- Laffel, G. e Blumenthal, D. (1989). The case for using industrial quality management science in health care organizations. *Jama*, 262(20):2869–2873.
- Li, X., Beullens, P., Jones, D., e Tamiz, M. (2009). An integrated queueing and multi-objective bed allocation model with application to a hospital in china. *Journal of the Operational Research Society*, 60(3):330–338.
- Mandelbaum, A., Momčilović, P., e Tseytlin, Y. (2012). On fair routing from emergency departments to hospital wards: Qed queues with heterogeneous servers. *Management Science*, 58(7): 1273–1291.
- Oliveira, R. B. O. d. e Rocha, F. G. C. (2020). Controle dinâmico de carga baseado em política de limiar aplicado a um sistema perinatal. *Anais do LII Simpósio Brasileiro de Pesquisa Operacional*, 2020.
- Pehlivan, C., Augusto, V., e Xie, X. (2013). Admission control in a pure loss healthcare network: Mdp and des approach. *Proceedings of the 2013 Winter Simulation Conference*, 32:5–44.
- Pehlivan, C. (2014). *Design and flow control of stochastic health care networks without waiting rooms: A perinatal application*. PhD thesis, Ecole Nationale Supérieure des Mines de Saint-Etienne.
- Pehlivan, C., Augusto, V., Xie, X., e Crenn-Hebert, C. (2012). Multi-period capacity planning for maternity facilities in a perinatal network: A queueing and optimization approach. In *2012 IEEE International Conference on Automation Science and Engineering (CASE)*, p. 137–142. IEEE.

- Tseytlin, Y. (2009). *Queueing systems with heterogeneous servers: On fair routing of patients in emergency departments*. PhD thesis, Technion-Israel Institute of Technology, Faculty of Industrial and
- Véricourt, F. d. e Jennings, O. B. (2011). Nurse staffing in medical units: A queueing perspective. *Operations Research*, 59(6):1320–1331.
- Wolff, R. W. (1989). *Stochastic modeling and the theory of queues*. Pearson College Division.
- Yankovic, N. e Green, L. V. (2011). Identifying good nursing levels: A queueing approach. *Operations research*, 59(4):942–955.