



UMA ANÁLISE DO SERVIÇO DE ATENDIMENTO MÓVEL DE URGÊNCIA NA ZONA OESTE DO RIO DE JANEIRO VIA SIMULAÇÃO A EVENTOS DISCRETOS

Musbah Koleilat

Universidade Federal do Rio de Janeiro – PEP – COPPE/UFRJ Av. Horácio Macedo, 2030 - Prédio do CT - Bloco F, sala 105 - Ilha do Fundão - RJ koleilat@po.coppe.ufrj.br

Ivail Muniz Junior

Universidade Federal do Rio de Janeiro – PEP – COPPE/UFRJ Av. Horácio Macedo, 2030 - Prédio do CT - Bloco F, sala 105 - Ilha do Fundão - RJ ivailmuniz@gmail.com

Ana Luiza Lima de Souza

CEFET/RJ – UnED Nova Iguaçu Estrada de Adrianópolis, 1317 – Santa Rita – Nova Iguaçu - RJ aluisouza@gmail.com

Mário Jorge Ferreira de Oliveira

Universidade Federal do Rio de Janeiro – PEP – COPPE/UFRJ Av. Horácio Macedo, 2030 - Prédio do CT - Bloco F, sala 105 - Ilha do Fundão - RJ mario_jo@pep.ufrj.br

RESUMO

Esse trabalho apresenta uma análise parcial do serviço de atendimento móvel de urgência na Área de Planejamento 4 (AP4) da Zona Oeste do Rio de Janeiro utilizando simulação a eventos discretos. Para isso, concebe-se um modelo de simulação a eventos discretos que busca representar o fluxo operacional do atendimento pré-hospitalar, desde a ocorrência de uma solicitação de resgate, até a entrega do paciente aos médicos no hospital. Tal modelo passa a incorporar novos elementos, levando agora em consideração a localização dos grupamentos da região, os recursos disponíveis nestes, e a demanda dos bairros. A partir dos dados coletados junto ao CBMERJ, dos resultados presentes na literatura, e das informações obtidas junto aos órgãos competentes, tem-se dois objetivos centrais: 1) determinar o tempo gasto neste atendimento pré-hospitalar; 2) identificar/avaliar as atividades críticas do processo, estimando-se os recursos necessários para a redução desse tempo de atendimento.

PALAVRAS CHAVE. Simulação, Atendimento de Urgência, Modelo Computacional. Área principal: Simulação.

ABSTRACT

This paper presents a partial analysis of the emergency mobile service of the Planning Area 4 (AP4) of the West Zone of Rio de Janeiro using discrete event simulation. To this end, we conceive a discrete event simulation model, which seeks to represent the operational flow of prehospital care, since the occurrence of a redemption request until the delivery of the patients to doctors at the hospital. This model is incorporating new elements, now taking into account the location of the clusters in the region, these resources available, and the demand of the neighborhoods. From the data collected from the CBMERJ, the results in the literature, and the information obtained from the competent bodies, has two main objectives: 1) to determine the time spent in pre-hospital care, 2) identify / evaluate activities critical process, estimating the resources required for the reduction of the time of service.

KEYWORDS. Simulation. Emergency Services. Computer Model. Main area: Simulation.



1. Introdução

A acessibilidade aos hospitais públicos, principalmente o acesso aos serviços de urgência ou emergência, tem se demonstrado um problema complexo, em virtude do crescimento da população e da falta de uma estrutura adequada para ofertar serviços de saúde de qualidade à população. Tal complexidade deriva de uma série de fatores, dentre os quais a dificuldade no dimensionamento e na aplicação dos recursos empregados no atendimento pré-hospitalar realizado. Embora a idéia do atendimento fora da estrutura hospitalar, com o uso de ambulâncias com suporte médico tenha se originado no final do século XIX, em Nova Iorque, e daí tenha sido levada para a Europa, é da França a concepção de regulação médica das urgências que se utiliza no Brasil. Segundo o site oficial do serviço de atendimento móvel de urgência francês, o "SAMU DE FRANCE" começou na década de 60 e o conceito de regulação médica das urgências tem sua origem no atendimento pré-hospitalar francês (Garcia, 2006). O Brasil iniciou a prática de atendimento pré-hospitalar (APH) especializado a partir dos anos 80, sendo o primeiro em 1986, através da fundação do Grupo de Socorro de Emergência (GSE), com sede no Rio de Janeiro, e o segundo, em 2003, com a criação do Serviço de Atendimento Móvel de Urgência (SAMU).

O serviço do SAMU funciona 24 horas por dia com equipes de profissionais de saúde, como médicos, enfermeiros, auxiliares de enfermagem e socorristas que atendem às urgências de natureza traumática, clínica, pediátrica, cirúrgica, gineco-obstétrica e de saúde mental da população. O socorro é feito após chamada gratuita, feita para o telefone 192. A ligação é atendida por agentes de atendimento no Centro de Operações GSE/SAMU (COGS). A função básica de um Serviço de Atendimento Móvel de Urgência (SAMU) é responder de forma organizada, a fim de evitar o uso excessivo de recursos, a toda situação de urgência que necessite de meios médicos, desde o primeiro contato telefônico até a liberação das vítimas ou seus encaminhamentos hospitalares. O sistema deve determinar e desencadear a resposta mais adequada para o caso, assegurar a disponibilidade dos meios hospitalares, determinar o tipo de transporte exigido e preparar o acolhimento dos pacientes, conforma aponta Takeda, Widmer, e Morabito (2001). Para se analisar o atendimento pré-hospitalar é necessário uma representação de seu fluxo operacional, o qual envolve diversos aspectos, dentre eles: um estudo das demandas dos tipos de atendimentos de urgência e emergência; determinação dos recursos humanos, tais como atendentes, médicos reguladores, equipes médicas, etc.; dimensionamento de ambulâncias e material necessário para atender à região analisada; comunicação entre agentes para reduzir o tempo total do paciente no processo, entre outros.

O objetivo deste trabalho é desenvolver um modelo de simulação a eventos discretos que seja capaz de representar o fluxo operacional do atendimento pré-hospitalar, desde a ocorrência de uma solicitação de resgate até a liberação do paciente, na Área de Planejamento 4 (AP 4), que integra a Zona Oeste da cidade do Rio de Janeiro, para que se reduza o tempo desse atendimento. Além disso, deseja-se determinar o tempo necessário para que uma vítima receba atendimento médico no local da ocorrência, por uma equipe de socorro especializado, visando verificar a conformidade do serviço de urgência nessa região em comparação ao critério internacional de 8 minutos, exposto por Junqueira e Dias (2012).

O trabalho se encontra dividido em seis sessões. Após a introdução, a segunda sessão aborda os aspectos de interesse do sistema real, para formular o modelo conceitual. Na terceira sessão foi feita a tradução do modelo conceitual para o software de simulação, que em seguida é utilizado para se realizarem os experimentos contidos na quarta sessão. Finalmente, são discutidos os resultados e feitas conclusões, na quinta e sexta sessões.

2. Simulação e Modelo Conceitual

A simulação é uma ferramenta utilizada para analisar o desempenho de um sistema por meio de uma modelagem, que é desenvolvida a partir de elementos conhecidos. O modelo constrói os componentes do sistema e é capaz de realizar experimentos que apoiam o entendimento do sistema real em determinadas condições. Segundo Bateman e Bowden (2002), a simulação é um processo de experimentação a partir de um modelo do sistema real para definir a maneira como este responderá a modificações em sua estrutura e ambiente. Para Banks (1984)



um modelo de simulação deve considerar os seguintes princípios no desenvolvimento de um experimento: o modelo deve ser o mais simples possível, embora precise ser completo nas questões relevantes, sendo atualizável e adaptativo, possuindo contribuições/recomendações feitas pelo usuário; precisa ser evolucionário de modo a ser desenvolvido em um baixo grau de complexidade, além de ser capaz de agrupar outras características caso necessário; possuir objetivos e finalidades bem determinados e direcionados.

Na construção de um modelo de simulação, são utilizadas como fundamentos algumas recomendações e técnicas, que são bastante difundidas pela literatura especializada no assunto. Baseado em Chwif e Medina (2007), desenvolveu-se o estudo elaborando um modelo conceitual do sistema, de forma a vislumbrarem-se os processos existentes no sistema real. Após serem mapeados os processos, pode ser iniciada a montagem (programação) do modelo computacional, utilizado para gerar os resultados da simulação. Depois, executam-se experimentos para verificar as alternativas que proporcionem os resultados esperados. Portanto, optou-se por utilizar a principio a linguagem natural para descrever as atividades que compõe o sistema, de modo que se pudesse identificar e definir o conjunto de elementos relacionados entre si, o meio circundante, e os limites considerados. Elaborou-se um modelo conceitual, onde foi utilizada a técnica de Diagrama de Ciclo de Atividades (Activity Cicle Diagram, ACD), obtendo as interações entre entidades, atividades e filas, tendo um entendimento mais amplo do sistema e dos relacionamentos entre os processos. A seguir descreveram-se mais detalhadamente todas as atividades pertencentes ao modelo desenvolvido nesse estudo, considerando as diferentes etapas do fluxo operacional do atendimento pré-hospitalar, que foram elaboradas utilizando informações obtidas por entrevista com profissionais do sistema GSE/SAMU, alocados no Centro de Operações GSE/SAMU (COGS) no ano de 2012.

Atendimento da Chamada (TARM): O agente de atendimento recebe a chamada inicial via telefone 192, faz o registro dos principais dados do evento via sistema informatizado, identifica o risco e tipo de evento, e direciona a chamada para a atividade seguinte, de acordo com a sua classificação de risco e local da ocorrência. O agente poderá: finalizar a chamada em caso de eventos diversos que não necessitem de tratativa (evento verde); direcionar a chamada diretamente para o despacho de frota (DESPURG) em caso de eventos de grande urgência (evento vermelho); ou ainda, caso não consiga identificar com precisão o tipo de atendimento necessário, poderá transferir a chamada para o médico regulador primário.

Regulação de Evento (MR 1): O médico regulador recebe a ligação 192 transferida pela TARM ou via DESPURG (Despacho de Urgência), confirma ou não a estratificação de risco feita pelo TARM. Recebida a ligação, ele pode fornecer orientações pelo telefone, muitas vezes solucionando a situação de urgência sem o envio de recursos (evento verde). Quando indicado, solicita o envio emergencial da viatura acionando o Despacho de Frota (evento amarelo). Pode fornecer orientações para urgências, mesmo após ter sido acionado o DESPURG.

Solicitação e Localização das Ambulâncias (DESPURG ou Despacho de Frota): No DESPURG ou Despacho de Frota, os Operadores de Frota, profissionais encarregados de acionar a ambulância, recebem a solicitação do TARM ou MR 1 através de sistema informatizado, e localiza a ambulância disponível mais próxima da ocorrência. Ao localizar uma ambulância disponível nos grupamentos próximos ou imediações do evento, ocorre a troca de informações entre o operador de frotas e a equipe da ambulância sobre a ocorrência, que é feita em sua quase totalidade com a utilização de rádio comunicador. Ao receber as informações sobre a localização da ocorrência, o motorista da ambulância define o percurso a ser seguido, baseado em sua experiência, e com a ambulância já em movimento, é passado o restante das informações, como por exemplo, o estado em que a vítima supostamente se encontra, entre outras. A solicitação é feita pelo DESPURG quando o agente de atendimento (TARM) identifica que uma chamada é um evento do tipo vermelho, exigindo que uma ambulância seja despachada imediatamente. Nesse caso, o tempo de despacho padrão é de no máximo 3 minutos. No caso de evento amarelo, o médico regulador é quem faz a solicitação de envio da ambulância, acionando o Despacho de Frota. Neste caso a ambulância pode ser enviada em um prazo máximo de 4 horas. Caso existam várias ocorrências ao mesmo tempo, existe a possibilidade de formação de fila nesta atividade.



Chegada ao local e atendimento inicial à vítima: Na chegada ao local são feitas as atividades de estacionamento da ambulância, transporte de equipamentos e acesso a vítima, que pode estar inacessível. Ao acessar-se a vítima, iniciam-se os procedimentos de socorro no local.

Transporte da vítima até o hospital informado (se necessário): Definido o local de destino, normalmente já conhecido pela equipe da ambulância que atua na região, a equipe inicia o transporte. O tempo de realização desta atividade depende principalmente da distância a ser percorrida e da fluidez do trafego para a ambulância.

Chegada ao hospital e passagem da vítima para a equipe médica: Ao chegar ao hospital é feita a remoção imediata da vítima da viatura, e a transferência da vítima aos cuidados dos médicos do hospital, com as devidas informações sobre a ocorrência, para que sejam iniciados os primeiros procedimentos, tão rápidos quanto se façam necessários. Em seguida, o oficial enfermeiro ou médico realiza a conclusão do preenchimento de formulário de registro do atendimento. Após a entrega do formulário ao hospital, a ambulância aguarda a liberação.

Liberação da ambulância: A liberação deveria ser automática, uma vez que a vítima foi entregue com explicações e relatório devidamente preenchido. Existe um protocolo de trabalho, no qual os hospitais não podem se recusar a receber os pacientes trazidos pelos bombeiros, mesmo que não haja vaga disponível, que visa economizar tempo ao realizar essa transferência. Na prática, como os hospitais possuem normalmente seus leitos ocupados, os pacientes ficam um longo período aguardando na maca dos bombeiros, o que gera atraso no ciclo do processo.

Preparo da ambulância: Uma ambulância só está pronta para atender uma nova ocorrência quando todo o seu material estiver adequadamente acondicionado e desinfetado. Depois de feito o atendimento da vítima, caso não seja necessário uma grande limpeza, a desinfecção é feita na própria viatura, para tornar mais rápida a preparação, pois pode ocorrer uma solicitação de atendimento no percurso de retorno a base.

Retorno à base: Ao iniciar o retorno à base, caso a preparação da ambulância seja concluída durante o percurso, ela estará disponível para novo atendimento, mesmo antes de chegar a base. Caso contrário, só estará disponível após retornar a base e concluir o preparo.

A Figura 1 abaixo apresenta o modelo ACD, baseado nas informações descritas:

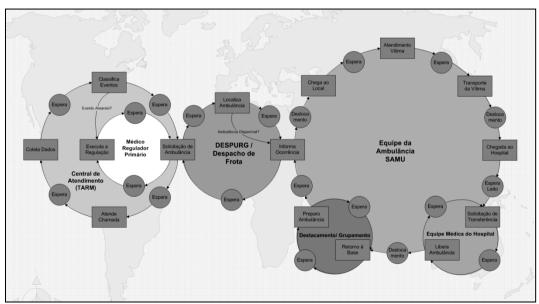


Figura 1 - Diagrama do Ciclo de Atividades do Atendimento de Urgência do SAMU/RJ.

Após a elaboração do modelo, possui-se um melhor entendimento sobre a dinâmica do sistema e quais atividades possuem maior impacto na parte do sistema que se deseja analisar. Esse modelo facilita substancialmente a tarefa de modelagem computacional, ao lançar luz sobre os aspectos principais.



3. Modelagem Computacional

Um sistema médico de emergência pode ser concebido como uma fila única operando sob um sistema de fila com múltiplos servidores subordinados a um controle central. A distribuição dos tempos de serviço é uma função da quantidade de ambulâncias que estão ocupadas em um determinado instante t caracterizado pela chegada de uma chamada. Além disso, a distribuição do tempo de atendimento não é a mesma para cada servidor, exceto para o caso em que todas as ambulâncias estão localizadas na mesma estação. Estas características do processo de atendimento médico de urgência dificultam imensamente a modelagem analítica deste sistema de filas, uma vez que a localização dos servidores e a distribuição espacial dos incidentes em cada localidade influenciam o tempo de deslocamento, que é um componente significativo do tempo de serviço, conforme apontado por Fitzsimmons (1971). A simulação, portanto, foi escolhida nesse estudo por se mostrar uma ferramenta adequada para lidar com incertezas associadas ao processo de resgate, às características estocásticas das atividades, além de permitir uma análise da disponibilidade de recursos e os impactos no tempo das atividades envolvidas.

A partir do modelo conceitual desenvolveu-se um modelo computacional, utilizando o software SIMUL8©, que é um simulador comercial implementado em Visual Basic®. Este estudo adotou a estratégia de programar um modelo "simplificado", conforme ilustrado na Figura 2, que depois será controladamente "complicado", conforme os testes de verificação e validação sejam bem sucedidos, baseados nos 5 princípios de Pidd (1996).

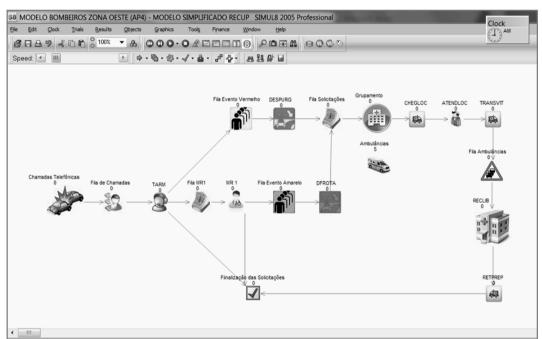


Figura 2 - Modelo Computacional Simplificado

Resumimos na Tabela 1 as atividades, parâmetros e fontes utilizados na confecção tanto do modelo simplificado, quanto do modelo completo:

Atividades:	Parâmetro:	Fonte(s):
Chamadas Telefônicas	Distribuição Exponencial (EXP[2])	Oliveira (2012)
TARM	Distribuição Triangular (TR[1,5;2;2,5])	Oliveira (2012)
MR1	Fixo (30 min)	Oliveira (2012)
DFROTA e DESPURG	Distribuição Exponencial (EXP[3,7]+1)	CBMERJ (2013)
CHEGLOC, TRANSVIT,	Tabela 4	Google (2012)
RETPREP.	1 abcia +	G00gic (2012)
ATENDLOC	Distribuição Weibull (WEIB[26,3;1,3] + 2,5)	CBMERJ (2013)
RECLIB	Distribuição Weibull (WEIB[45.4,1.3] + 5).	CBMERJ (2013)

Tabela 1 - Parâmetros utilizados no modelo.



Para o bloco "Chamadas Telefônicas", além da inserção do parâmetro exponencial apresentado acima, foi necessária uma configuração adicional. Cada chamada possui uma probabilidade de ser classificada em evento vermelho (70%), amarelo (17%) ou verde (13%), conforme mencionado na descrição das atividades, utilizando dados de Garcia (2006). Para ajustar essa característica no modelo computacional é necessária a configuração de um rótulo (label) no bloco "Chamadas Telefônicas", de modo que cada chamada é classificada em evento vermelho, amarelo ou verde, seguindo as probabilidades mencionadas. Após a inserção das distribuições e configurações das filas, entidades e atividades, o modelo pode ser executado. Para maiores detalhes sobre programação computacional utilizando o SIMUL8 consultar Concannon (2007). O modelo computacional se diferenciou visualmente do modelo conceitual apenas pela inserção dos blocos de "filas", o que permite a análise de resultados entre esses elementos.

Iniciando a fase de verificação e validação do modelo, que consiste, de acordo com Chwif e Medina (2007), em se retirar *bugs* do modelo e verificar se o comportamento do modelo está de acordo com o modelo conceitual, foi verificada a configuração de cada bloco. Após está etapa, foi necessária a definição do tempo de simulação e número de replicações para que se pudesse avançar para a fase de validação, ou seja, se os resultados do modelo são condizentes com o mundo "real". Podemos facilmente obter uma boa estimativa para o tempo de simulação, baseados no tempo de *warm-up*, que é tempo necessário para que as condições iniciais do sistema não afetem os resultados que serão obtidos do modelo. Usualmente, procuramos trabalhar com tempos de simulação superiores a no mínimo, 10 vezes o tempo de *warm-up* (regra empírica). Utilizando a técnica da "observação", conforme Chwif e Medina (2007), determinou-se os parâmetros para a obtenção do tempo de aquecimento e warm-up, reduzindo os efeitos das condições iniciais da partida do sistema, conforme ilustra a Figura 3.

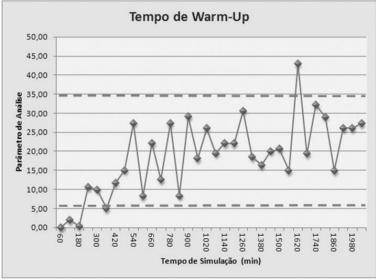


Figura 3 – Gráfico de definição do tempo de warm-up.

Utiliza-se no presente modelo o tempo de aquecimento (t_w) de 2.000 minutos, e em consequência o tempo de simulação (t_s) de 15 dias (21.600 minutos). A partir dos tempos de *warm-up* e de simulação, iniciamos o cálculo do número de replicações, baseados em Chwif e Medina (2007), conforme as equações (2) e (3). Para o cálculo do número de replicações realizase a simulação do parâmetro escolhido, várias vezes, com o intuito de calcular a precisão, conforme apresentado na Tabela 2.

Replicação (n)	1	2	3	4	5	6	7	8	9	10
Parâmetro	33,54	39,82	28,27	32,76	26,28	25,20	29,30	30,97	23,99	38,19

Tabela 2 - Valores obtidos por simulação do parâmetro escolhido.



Iniciamos então nosso cálculo com 10 replicações (n = 10) e os resultados da Tabela 2, para o parâmetro simulado e seu desvio padrão (s).

$$h = t_{n-1,\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} \tag{2}$$

$$n^* = n \cdot \left(\frac{h}{h^*}\right)^2 \tag{3}$$

Assim, utilizando os dados da Tabela 2, e aplicando a equação (2) acima, temos que a precisão (h) para este parâmetro é de 3,8 minutos, com um nível de confiança de 95%. A partir desses valores e utilizando a equação (3), calcula-se o número de replicações necessárias (n^*) para que com um nível de confiança de 95% e uma amostra com n = 10, obtenha-se uma precisão de 2 minutos (h^* = 2). Portanto, será utilizado um número de replicações igual a 37 na simulação desse modelo. Obtidos os parâmetros que faltavam para iniciar as simulações e consequentes experimentações, avançou-se com a fase de validação do modelo, utilizando a técnica de validação denominada análise de sensibilidade, que consiste em determinar a influência de alterações dos parâmetros de entrada nos resultados obtidos a partir do modelo. Foram definidas quatro verificações para a análise de sensibilidade, que são descritas e avaliadas abaixo, a partir dos resultados da Tabela 3:

Disease	D-u2tu-	Verific	ação 1	Verific	ação 2	Verific	ação 3	Verificação 4		
Bloco	Parâmetro	Mod. Orig.	Mod. Alt.	Mod. Orig.	Mod. Alt.	Mod. Orig.	Mod. Alt.	Mod. Orig.	Mod. Alt.	
	Tamanho médio da fila	0	14,19	-	-	-	-	-	-	
Fila de	Tamanho máximo da fila	1	30	-	-	-	-	-	-	
Chamadas	Tempo médio na fila (min)	0	29,5	-	-	-	-	-	-	
	Tempo máximo na fila (min)	0	59,0	-	-	-	-	-	-	
	Tamanho médio da fila	-	-	0,66	1,62	0,66	0,92	-	-	
Fila do MR1	Tamanho máximo da fila	-	-	2,26	4,17	2,26	2,88	-	-	
FIIa do IVIKI	Tempo médio na fila (min)	-	-	15,9	28,13	15,98	16,0	-	-	
	Tempo máximo na fila (min)	-	-	35,74	58,76	35,74	36,5	-	-	
	Tamanho médio da fila	-	-	0,01	0,03	0,01	0,01	-	-	
Fila Evento	Tamanho máximo da fila	-	-	1	1,10	1	1,01	-	-	
Vermelho	Tempo médio na fila (min)	-	-	0,31	0,70	0,31	0,22	-	-	
	Tempo máximo na fila (min)	-	-	1,49	3,40	1,49	1,32	-	-	
	Tamanho médio da fila	-	-	0,47	1,35	0,47	1,66	-	-	
Fila	Tamanho máximo da fila	-	-	2,31	4,76	2,31	5,61	-	-	
Solicitações	Tempo médio na fila (min)	-	-	5,13	11,5	5,13	13,6	-	-	
	Tempo máximo na fila (min)	-	-	23,02	48,6	23,02	56,4	-	-	
	Tamanho médio da fila	-	-	-	-	0,39	0,61	0,39	0,85	
Fila	Tamanho máximo da fila	-	-	-	-	1,94	2,58	1,94	2,75	
Ambulâncias	Tempo médio na fila (min)	-	-	-	-	13,9	19,7	13,9	23,0	
	Tempo máximo na fila (min)	-	-	-	-	28,2	39,7	28,2	46,5	

Tabela 3 - Tabela com resultados da validação do modelo.

Verificação 1 - Aumento em 50% no número de ligações: Como resultados esperados, é previsto um aumento na fila de chamadas, pois a atividade TARM não terá capacidade de processar essas chamadas imediatamente. O resto do sistema deve permanecer inalterado, pois quem determina o ritmo no qual as chamadas avançam no sistema é o TARM. Os outros resultados do modelo permaneceram inalterados. Os resultados apresentados na Tabela 3 confirmaram a previsão inicial, pois efetivamente se formou uma fila antes do TARM.

Verificação 2 - Aumento em 50% no número de ligações e na capacidade de atendimento no TARM: Neste caso, há redução na fila de chamadas, pois o TARM agora passa a ser eficiente no atendimento as chamadas. Os gargalos do sistema passam a ser o DESPURG e o MR1, com a formação de filas antes dessas atividades. O restante do sistema permanece inalterado, pois quem determina o ritmo no qual as chamadas e solicitações avançam no sistema são o DESPURG e MR1. Os dados confirmam a previsão inicial, pois o tempo nas filas do "MR1" e "Evento Vermelho" aumentaram. Além disso, constatou-se aumento na utilização das ambulâncias em 10%, e aumento considerável no tempo da "Fila Solicitações".

Verificação 3 - Aumento em 50% na capacidade de atendimento, tanto no MR1 quanto



no DESPURG: O aumento da eficiência nos atuais gargalos do sistema (DESPURG e o MR1) irá reduzir suas respectivas filas, "Fila Evento Vermelho" e "fila MR1", voltando aos padrões normais, e aumentando ainda mais a "Fila Solicitações" de ambulâncias, assim como a utilização das ambulâncias em si. A "Fila Ambulâncias" no hospital (RECLIB), também deve ter ligeiro aumento, pois serão enviadas mais ambulâncias ao hospital. Após verificação, constata-se que todas as previsões feitas acima se confirmaram no modelo computacional.

Verificação 4 - Aumento em 50% na capacidade de atendimento da equipe da ambulância no local do acidente: O aumento da eficiência no atendimento no local aumentará ainda mais a demanda por atendimento no hospital (RECLIB). Consequentemente, a "Fila Ambulâncias" no hospital esperando pelo atendimento também aumentará, visto que a taxa de chegada aumenta e a taxa de liberação no hospital foi mantida. Os dados obtidos nesse cenário confirmaram as previsões, indicando que o modelo é sensível aos parâmetros alterados durante os experimentos, o que permitiu finalizar a etapa de validação do modelo.

Após a verificação e validação do modelo simplificado, que tornou o modelo de simulação operacional pronto para ser utilizado, foi possível dar mais um passo na direção da construção de um modelo que expressasse com mais detalhes a realidade. Nessa última fase, considera-se os 4 grupamentos e os 19 bairros existentes na região de estudo, ampliando o modelo em tamanho e complexidade. Para isso, considera-se três blocos de atividade referentes aos grupamentos que existem na região (12GBM, GBS/GMAR e PABM), e 19 blocos de atividade CHEGLOC relacionando o tempo de trajeto do grupamento até o local do acidente, considerando cada um dos 19 bairros da AP4. Em seguida, configura-se 19 blocos de atividade ATENDLOC, um para cada bairro, que representam o atendimento das vítimas no local da ocorrência. Insere-se também 19 blocos de atividade TRANSVIT representando o tempo de transporte da vítima socorrida ao hospital de emergência de destino. O tempo de deslocamento do grupamento ao local da ocorrência, e da ocorrência até o hospital de destino, foi estimado utilizando-se a ferramenta Google Maps® (2012). Esses tempos são apresentados na Tabela 4:

	T:	abela	a de	Tem	pos	(mir	ı) en	tre o	s Ba	airro	s da	Zon	а Ос	este	(AP	4)				
	De/	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133
	Para	JP	AN	GA	$^{\rm CD}$	CC	FR	PC	TQ	TN	PS	VV	JO	IT	BT	CM	VP	VG	RB	GR
	115	0	9	9	9	19	7	11	18	15	18	24	29	26	15	28	31	35	26	35
	116	9	0	7	11	20	3	14	20	17	20	26	26	22	17	30	32	37	27	37
	117	8	6	0	9	18	5	13	18	16	19	25	31	28	15	28	30	35	25	35
	118	12	13	11	0	12	12	7	12	10	13	19	27	24	13	22	27	32	27	33
	119	21	22	20	12	0	21	13	12	12	15	21	34	32	20	14	19	25	36	32
	120	-6	3	5	9	18	0	-11	18	15	18	23	27	23	15	28	30	35	25	34
	121	11	9	12	7	15	9	0	13	5	9	14	27	25	13	25	30	35	23	33
	122	22	20	20	11	12	20	12	0	10	12	18	37	34	23	22	27	32	34	40
12º GBM	123	16	15	18	11	14	15	6	10	0	6	11	33	30	19	24	29	34	29	39
	124	20	18	23	13	16	18	10	12	7	0	7	37	34	23	26	31	36	29	44
•	125	24	23	27	17	20	22	14	17	11	7	0	41	38	27	30	35	41	31	48
	126	24	26	22	23	28	23	25	32	29	32	37	0	7	13	36	32	36	26	36
	127	21	23	20	21	26	20	22	30	26	29	35	6	0	11	33	29	33	27	33
GBS/GMAR	128	16	17	15	16	21	16	18	25	21	25	30	21	18	0	28	18	23	29	23
	129	28	29	27	24	20	28	29	29	28	31	36	40	37	26	0	18	23	42	33
	130	30	31	29	25	21	30	30	30	29	32	37	36	34	25	24	0	8	44	24
	131	35	36	34	30	26	35	35	35	34	37	42	41	38	29	29	8	0	50	23
PABM-01	132	28	26	28	28	34	26	26	36	28	32	36	25	27	29	41	45	49	0	49
	133	35	37	34	35	36	35	37	45	40	44	49	38	36	27	41	27	22	47	0

Tabela 4 - Tabela de tempos entre bairros.

A distribuição das chamadas pelos bairros foi estimada, considerando-a diretamente proporcional às populações de cada bairro, cujos números foram obtidos de Prefeitura do Municipio do Rio de Janeiro (2013). Para configurar cada um dos blocos relativos aos bairros, foi necessária a criação de outros dois rótulos, o rótulo "Grupamento" e o rótulo "Bairros", para adequado encaminhamento das solicitações.

4. Experimentação

Iniciando então a análise global do sistema, são examinadas as atividades críticas do sistema, ou seja, aquelas que estão gerando filas, visando verificar se os dados de entrada e parâmetros do modelo estão consistentes. No caso de inconsistência(s), deve-se atualizar esses



parâmetros, pois provavelmente esses dados não são suficientemente precisos. A partir do momento, em que os dados foram ajustados e o modelo está fornecendo resultados confiáveis, procede-se com os experimentos. Na etapa de experimentos, através de modificações envolvendo a quantidade de recursos e/ou parâmetros de processamento de tarefas, encontram-se novos valores de tempos de filas, de atividade e de ciclos do processo, conforme os recursos ou atividades vão sendo balanceados. Como consequência, objetiva-se reduzir o tempo máximo do ciclo do serviço de emergência. A Tabela 5 mostra os resultados do modelo com os parâmetros definidos inicialmente e adequados a cada novo experimento. As células em cinza possuem parâmetros que estão com valores que necessitam ser adequados.

Observou-se nas linhas dos itens 2, 5 e 10 com relação à coluna de resultados iniciais, que as filas possuem tempo de espera excessivo. Vemos também nos itens 3, 6 e 13 que as atividades e o recurso estão trabalhando quase 100% do tempo, condição limite inadequada. A primeira análise é feita na "Fila de Chamadas", que é o primeiro bloco crítico. Observando o modelo conceitual, percebe-se que essa fila esta interagindo diretamente com os blocos "Chamadas Telefônicas" e "TARM". Como o bloco "Chamadas Telefônicas" é um bloco de entrada de dados no sistema, ou seja, a porta de entrada com as solicitações do mundo real, ele não pode ser modificado, pois não se tem controle sobre a demanda da população. Por outro lado, o bloco "TARM" pode ser modificado, isto é, pode-se torná-lo mais eficiente, ao efetuar ações que melhorem a eficiência dessa entidade. Dentre as ações, o treinamento dos atendentes, simplificação do sistema de entrada de dados, redução do número de etapas no processo de transmissão de informações, e outras, podem ajudar o sistema a reduzir o tempo de atendimento nesta etapa. Iniciam-se as primeiras modificações, ao realizar os experimentos descritos.

Experimento 1: Supõem-se que podem ser efetuadas ações que consigam melhorar em 50% a eficiência da atividade "TARM", reduzindo a média e os extremos da distribuição triangular para 1.33, 1.00 e 1.66, respectivamente. Os resultados podem ser observados na coluna relativa ao experimento 1 da Tabela 5. Após a simulação com os novos parâmetros, observa-se uma redução considerável no tempo de espera da "Fila de Chamadas" de 104 minutos para 1,3 minutos no tempo médio de espera. Este valor está muito mais próximo da realidade, pois considera que não haveria espera de mais de 2 minutos para falar com os atendentes. Para tempos de espera maiores, o índice de desistência tende a crescer consideravelmente, afetando negativamente o atendimento pré-hospitalar. Outro ponto de melhora foi o percentual de utilização do TARM, que passou de 100% para 66,3%, configurando uma situação de trabalho mais condizente com a realidade. Voltando ao nosso modelo, observa-se agora que o bloco crítico passou a ser a "Fila MR1", que é o bloco a ser analisado no próximo experimento.

Experimento 2: Ao analisar o bloco "Fila MR1", que está conectado aos blocos "TARM" e "MR1", verifica-se que o número de casos enviados pelo TARM está acima da capacidade de atendimento e processamento do MR1. Assim, as chamadas a serem processadas estão se acumulando, gerando uma fila indesejável nessa parte do processo. Para reduzir a fila, simulam-se aumentos graduais do nível de eficiência do bloco "MR1", partindo de um aumento de 50%, e em seguida testando outras melhoras (100%, 200% e 300%), de modo que se obtenha um tempo médio de espera na fila MR1 mais factível. Os resultados podem ser analisados nas colunas relativas ao experimento 2.1 a 2.4.

Experimento 3: Ao observar que a "Fila do 12 GBM" é precedida pelo DESPURG, que envia as solicitações de ambulâncias para o 12 GBM, constata-se que a fila está sendo formada, pois o 12 GBM não está conseguindo processar essas solicitações no tempo necessário. Para aumentar a velocidade no atendimento, pode-se aumentar o número de ambulâncias disponíveis no 12 GBM. Aqui surge um ponto crítico, pois a taxa de utilização dessas ambulâncias é de 100%, ou seja, estão totalmente saturadas. Assim, se forem aumentados o número de ambulâncias, haverá acréscimo no número de atendimentos às vítimas. É claro, que este aumento de ambulâncias precisa ser gradual e até certo limite. Para ilustrar esse conceito, considera-se por suposição, deficiência do quantitativo de leitos nos hospitais. Neste caso, pode ocorrer de cada ambulância adicional incluída no sistema ser ocupada por tempo indeterminado pelo paciente, pois não existe leito disponível para o paciente ao chegar ao hospital. Isso faz com que a própria



ambulância seja transformada em leito, saturando rapidamente cada ambulância incluída no sistema. Portanto, caso se inclua ambulâncias no sistema, e não ocorra redução da fila de solicitações e da utilização das ambulâncias do grupamento, deve-se interromper a inclusão de ambulâncias e identificar outras partes do processo que podem ser gargalos do sistema, como os hospitais, por exemplo.

Item	Bloco	Tipo	Índica-dores	UM	Results Iniciais	EXP. 1 TARM 50%	EXP. 2.1 MR1 50%	EXP. 2.2 MR1 100%	EXP. 2.3 MR1 200%	EXP. 2.4 MR1 300%	EXP. 3.1 Remane- jamentos	EXP. 3.2 Ambul. adicional (12 GBM)
1	Chamadas Telefônicas	Entrada	Chamadas Recebidas	unid	10.761	10.761	10.791	10.761	10.761	10.761	10.761	10.761
2	Fila de Chamadas	Fila	Tempo Médio Tempo Máximo	min min	104,33 247,31	1.32 ✓	1,34 13,95	1,32 13,95	1,32 13,95	1,32 13,95	1,32 13,95	1,32 13,95
3	TARM	Atividade	Trabalhando	%	99,31	66.3 ✓	66,3	66,4	66,4	66,4	66,4	66,4
4	Fila Evento	Fila	Em Espera Tempo Médio	min	0,69 1,24	33,7 1,51	33,7 1,51	33,6 1,51	33,6 1,51	33,6 1,51	33,6 1,51	33,6 1,51
5	Vermelho Fila MR1	Fila	Tempo Máximo Tempo Médio	min min	27,95 7687	29,58 7755	29,58 5252	29,58 2760	29,58	29,58 6.45 ✓	29,58 6,45	29,58 6,45
6	MR1	Atividade	Tempo Máximo Trabalhando	min %	14218 100	14321 100	9740	5138 100	153 85,2	54.92 ✓ 63.9 ✓	54,92 63,9	54,92 63,9
7	Fila Evento	Fila	Em Espera Tempo Médio	% min	0	0	0 0,01	0,03	14,8 0,1	36,1 0,18	36,1 0,18	36,1 0,18
8	Amarelo DESPURG	Atividade	Tempo Máximo Trabalhando	min %	0,03 30,4	0,03 30,5	1,69 30,5	6,42 30,5	12,21 30,5	14,31 30,5	14,31 30,5	14,31 30,5
9	DFROTA	Atividade	Em Espera Trabalhando	%	69,6 5 95	69,5 5	69,5 7,5	69,5 10	69,5 12,7	69,5 12,7	69,5 12,7	69,5 12,7
10	Fila do 12	Fila	Em Espera Tempo Médio	% min	934	95 980	92,5 1609	90 2260	87,3 2916	87,3 2927	87,3 61	87,3 8.01 ✓
11	GBM Fila do	Fila	Tempo Máximo Tempo Médio	min min	1845 0,69	1915 0,78	3065 1,03	4278 1,4	5486 1,77	5492 1,87	287 8,83	96.09 √ 8,83
12	GBS/GMAR Fila do PABM	Fila	Tempo Máximo Tempo Médio	min min	37,8 0	39,49 0	45,22 0	49,5 0	53,72 0	53,28 0	111,79 0	111,79 0
	Ambulâncias		Tempo Máximo Utilização	min %	99,6	99,6	99,9	0,08 99,99	0,57 100	0,36 100	0,36 87,6	0,36 65.91 ✓
13	12 GBM Ambulâncias	Recurso	Uso Médio Utilização	unid %	1,99 37	1,99 37,2	2	2 42,5	2 45,3	2 45,3	2,63 60,13	2,64 60,05
14	GBS/GMAR	Recurso	Uso Médio	unid	1,48	1,49	1,65	1,7	1,8	1,8	1,81	1,8
15	Ambulâncias PABM	Recurso	Utilização Uso Médio	% unid	0	0	0	0	0	0	0	0
16	Fila Ambulâncias	Fila	Tempo Médio Tempo Máximo	min min	1,77 19,33	1,77 18,96	1,88 19,1	2 19,5	2,13 20,47	2,12 21,58	2,54 21,74	2,81 23,66
17	RECLIB	Atividade	Trabalhando Em Espera	%	45,5 54,5	45,6 54,4	47,5 52,5	49,1 50,9	50,8 49,2	50,8 49,2	57,73 42,27	57,76 42,24
18	Finalização Solicitações	Saída	Tempo Máximo do Ciclo de Atividades	min	14.918 (248 h)	14.906 (248 h)	11.459 (190 h)	8.494 (141 h)	5.565 (93 h)	5.551 (93 h)	369 (6 h)	203 (3,4 h)

Tabela 5 - Resultados finais do modelo de simulação.

Como o aumento ou remanejamento do número de ambulâncias é um parâmetro que está sob controle do serviço de emergência, inicia-se o experimento pelo remanejamento das ambulâncias entre os grupamentos, visto que a taxa de utilização das 4 ambulâncias do grupamento GBS/GMAR é de 45,3%, bem menor que os 100% de utilização das 2 ambulâncias do 12 GBM. Assim, apresentam-se dois cenários adicionais que comporão o experimento 3: o primeiro, é transferir uma (1) ambulância do GBS/GMAR para o 12 GBM, ficando com 3 ambulâncias no 12 GBM e 3 ambulâncias no GBS/GMAR, onde os resultados são apresentados na coluna relativa ao experimento 3.1; o segundo, é melhorar ainda mais o serviço após o remanejamento, adquirindo outra ambulância, totalizando 4 no 12 GBM e 3 no GBS/GMAR, com resultados apresentados na coluna relativa ao experimento 3.2. Observa-se pelos resultados do remanejamento de ambulâncias (experimento 3.1), que houve redução significativa no tempo médio de espera das solicitações de ambulância, que era de 2.927 minutos e foi para 61 minutos. Esse valor reduziu ainda mais, chegando ao patamar de 8 minutos, ao se incrementar uma ambulância no 12 GBM (experimento 3.2). Portanto, constatamos que a alteração no número de ambulâncias se adequou bem ao sistema, não se saturando por completo os parâmetros modificados, e trazendo para um valor mais realista, a utilização das ambulâncias dos



grupamentos. Além disso, foi possível perceber grande melhora do tempo global do sistema com esta modificação, cujo valor máximo do ciclo de atividades passou de 93 h para 6 h, com o remanejamento das ambulâncias, e para 3,4 h com a modificação final.

5. Resultados

Neste ponto do estudo, pode-se obter resposta do modelo elaborado, relativo ao tempo necessário (tempo médio e máximo) para a chegada da ambulância no local da ocorrência, por uma equipe de atendimento médico de emergência. Para isso, executa-se o modelo já considerando os ajustes feitos ao longo dos experimentos, incluindo aqueles que exigem investimentos (aquisição de uma ambulância). A Tabela 6 abaixo resume os dados do modelo referente às atividades que vão desde a chamada telefônica até a chegada da ambulância no local da ocorrência, considerando a estrutura vigente, para um atendimento a evento ocorrido na AP4:

Bairros	Anil	G.Az.	C.D.	Curic.	Freg.	Pec.	Taq.	Tan.	P.S.	Val.	Cam.	Ita.	Joá	Barra	V.Pq.	V.Gd.	Jac.	Gru.	Rec.
Tempo Médio cheaada amb.	30,8	34,7	26,3	26,2	30,5	19,5	24,6	13,9	19,1	25,8	38,5	34,1	33,9	15,5	31,8	38,3	31,2	38,2	13,1
Tempo Máximo	64.2	84.6	81 9	64.0	86.4	59 4	72 5	58.8	60.4	51 9	82.5	92.2	48.6	82.4	73.8	67.8	88.6	118,0	58.3
chegada amb.	07,2	04,0	01,5	04,0	00,4	33,4	12,3	30,0	00,4	31,3	02,3	32,2	40,0	02,4	73,0	07,0	00,0	110,0	30,3

Tabela 6 - Simulação de tempos de chegada.

A partir dos resultados da Tabela 6, o tempo gasto por uma ambulância para chegar aos bairros da AP4 pode chegar a quase duas horas (118 minutos). Apesar dos tempos de atendimento nos bairros que possuem grupamentos localizados neles próprios, tais como Recreio (13,1 min), Tanque (13,9 min) e Barra (15,5 min) serem bem mais próximos do critério de 8 minutos, ainda existe uma diferença significativa, muitas vezes decorrente dos atrasos que acontecem antes mesmo da informação chegar ao grupamento de onde será despachada a ambulância. Verificou-se um consumo de tempo significativo no despacho, pois nessa etapa se perde em média 6,2 minutos, que foram contabilizados ao se observar os tempos consumidos nos blocos "Fila de Eventos Vermelhos", 1,5 minutos e o tempo de processamento do "DESPURG", que é de 4,7 minutos em se considerando a média da exponencial, vide Tabela 1. Com essas etapas já consumimos 77% do critério internacional de 8 minutos. Observa-se que para os bairros de Grumari, Vargem Grande e Camorim, onde todos eles possuem tempo de atendimento por ambulância acima de 35 minutos, o tempo necessário para a chegada das ambulâncias será elevado mesmo que se invista capital público para a compra de mais ambulâncias. Isso ocorre devido a distancia desses locais aos grupamentos existentes. Para que se obtenha redução de tempo para tais casos, se faz necessário a instalação de grupamentos e destacamentos mais próximos desses locais.

6. Conclusões

O estudo aqui apresentado teve como objetivo analisar o atendimento pré-hospitalar na Área de Planejamento 4, da cidade do Rio de Janeiro, o qual se configura como um problema complexo, pois envolve diversos agentes, cujas atividades nem sempre possuem dados disponíveis e confiáveis. A partir dos dados coletados, dos resultados presentes na literatura, e das informações obtidas junto aos órgãos competentes, foi elaborado um modelo de simulação a eventos discretos que se aproximasse da realidade encontrada nessa região. Construído o modelo, verificou-se a necessidade de adequação de alguns parâmetros coletados originalmente para uma resposta mais coerente do modelo em comparação com a "realidade". Por exemplo, ao observar o tempo de atendimento dos médicos reguladores, constata-se que para obter um tempo médio de espera da "Fila MR1" em torno de razoáveis 6 minutos, ao invés dos 7.000 minutos do modelo com dados iniciais, seria necessária melhora de eficiência do bloco "MR1" em 300%, ou seja, cair o tempo de atendimento que era de 30 minutos para 7,5 minutos. Após adequação desses parâmetros, identificou-se a necessidade de um aumento do número de ambulâncias no 12º GBM, localizado no Tanque, por ser o grupamento mais próximo da maioria dos bairros (11 ao todo). Tal demanda não é satisfatoriamente atendida com a quantidade atual de ambulâncias, e melhores resultados podem ser obtidos através de uma ação combinada entre remanejamento e aquisição de



ambulâncias. Com a passagem do número de ambulâncias de 2 para 4, o tempo médio da fila do 12° GBM passou de 2.297 minutos para aproximadamente 8 minutos. Ressalta-se que a ociosidade das ambulâncias observada no GMAR/GBS pode ser decorrente não somente da baixa demanda dos bairros, mas também pela falta de dados sobre os eventos na orla da praia.

Finalmente, atingiu-se o objetivo do trabalho, pois o modelo de simulação representa o fluxo operacional do atendimento pré-hospitalar, e serve como ferramenta para a realização de tantos experimentos quanto se queira fazer, a fim de se analisar o comportamento desse fluxo. Além disso, avaliou-se o sistema atual, com relação aos recursos disponíveis, concluindo que o critério internacional de 8 minutos não é factível utilizando-se os recursos considerados originalmente. Os problemas aqui identificados e as ações sugeridas constituem uma contribuição importante do trabalho, pois mostram pontos críticos do fluxo operacional desse atendimento na AP4, bem como algumas soluções que podem ser implementadas visando reduzir o tempo desse fluxo, e, consequentemente, aumentar as chances de recuperação e salvamento das pessoas em situação de emergência.

Referências:

Banks, J. e Carson, J. S. (1984). Discrete-Event System Simulation. N.J.: Prentice-Hall. Bateman, R. E. e Bowden, R. O. (2002). Simulação otimizando sistemas (2 ed.). SP: IMAM. Chwif, L. e Medina, A. C. (2007). Modelagem e Simulação de Eventos Discretos: Teoria e Aplicações. São Paulo: Editora do Autor.

Concannon, K. (2007). Simulation Modeling with SIMUL8. Visual Thinking International. Corpo de Bombeiros Militar do Estado do Rio de Janeiro (2013). SAMU CBMERJ. Acesso em 20 de 01 de 2013, disponível em http://www.samu.cbmerj.rj.gov.br/acompanhamento/ Fitzsimmons, J. (1971). An emergency medical system simulation model. Winter Simulation Conference. 5, pp. 18-25. New York: New York Proceedings.

Garcia, L. C. (2006). *Dimensionamento de Recursos de Atendimento Móvel de Urgência da Região Metropolitana II do Estado do Rio de Janeiro*. RJ: Dissertação Mestrado, COPPE/UFRJ. **Google** (2012). *Google Maps*. Acesso em 9 de Setembro de 2012, disponível em Google: http://maps.google.com.br/

Junqueira, F. e Dias, T. (4 de Junho de 2012). *Tempo de Espera pelo Samu na Zona Oeste é maior que a média*. Acesso: 22/04/2013, disponível: http://extra.globo.com/noticias/rio/tempode-espera-pelo-samu-na-zona-oeste-maior-que-media-5111145.html

Ministério da Saúde. (25 de 03 de 2010). *Portal da Saúde*. Acesso: 10 /01/2013, disponível: http://portal.saude.gov.br/portal/aplicacoes/noticias/default.cfm?pg=dspDetalheNoticia&id_area= 124&CO_NOTICIA=11168

Oliveira, D. G. (2012). Simulação do Fluxo Operacional do Serviço de Atendimento Pré-Hospitalar Realizado pelo GSE/SAMU no Municipio do Rio de Janeiro. RJ: Dissertação Mestrado - COPPE/UFRJ.

Pidd, M. (1996). Five Simple Principles of Modelling. *Proceedings of the Winter Simulation Conference*, pp. 721-728.

Prefeitura do Municipio do Rio de Janeiro (2013). *Bairros Cariocas*. Acesso em: 30 de Janeiro de 2013, disponível em: http://portalgeo.rio.rj.gov.br/bairroscariocas/index_bairro.htm **Silva, P.M.S.** (2010). *Análise do serviço de atendimento móvel de urgência (SAMU) de Belo Horizonte via simulação e otimização*. MG: Dissertação de Mestrado - UFMG.

Samu de France (2013). Portal. Acesso: 05/01/2013, disponível: http://www.samu-de-france.fr/fr/vie_samu/presentation

Takeda, R. A., Widmer, J. A. e Morabito, R. (2001). Uma Proposta Alternativa para Avaliação do Desempenho de Sistemas de Transporte Emergencial de Saúde Brasileiros. *Transportes*, 9 (2), pp. 9-27.