

# Análise Estocástica para Redução dos Custos de Contratação de *Call Center* da CPFL Energia<sup>\*</sup>

Erick Busato, Alexandre Aranã, Douglas Akassaka,  
Lídia Gusmão, Hugo Helito e Júlia Ramos

*Centro Analítico, CPFL Energia, Campinas-SP,  
(e-mail: eabusato@gmail.com, {alexandre, lguimaraes, douglasakassaka,  
hugoh, jbeatrizr}@cpfl.com.br).*

**Abstract:** This study aims at developing a predictive model to forecast the Call Center invoice calls volume in order to optimize the contracting of service volume. In doing so, it is possible to reduce operational costs, whilst offering a good quality service, and thus, mitigating regulatory risk. For a better model definition, besides service historical data and meteorological variables, a stochastic approach was carried out, considering a set of random events to project the volume of incoming calls. A stochastic projection for Average Talk Time (ATT) was also performed, in which Kernel marginal models and temporal correlation structure were applied to generate weekly scenarios on an hourly basis. Afterwards, a weekly projection model of 48 daily stations was made, using Kernel distribution and coupling functions, in which 2,000 scenarios were built with the due temporal correlation between the hour stations and days of the week, so that attendance indicators, given a number of attendants, were evaluated. Finally, a tool was implemented to dynamically calculate incall service volume, in which the user is able to adjust different parameters such as considered time period and number of attendants to obtain the aimed projection.

**Resumo:** O propósito deste estudo foi desenvolver uma análise sobre a volumetria de chamadas de *Call Center* de uma distribuidora e implementar um modelo preditivo retroalimentável para cálculo da demanda necessária. Foram levadas em conta as variáveis históricas do atendimento e a possibilidade de inclusão de variáveis meteorológicas para desenvolvimento de uma projeção estocástica do volume de chamadas recebidas. Foi realizada uma projeção estocástica para TMA (Tempo Médio de Atendimento), em que, com os modelos marginais *Kernel* e a estrutura de correlação temporal, foi obtido o modelo de simulação semanal de chamadas, gerando cenários em base intra horária. Também foi desenvolvido o modelo de projeção semanal de 48 postos diários (janelas de 30 minutos), em que foi utilizada a função de distribuição *Kernel* e função de acoplamento para análise de volumetria intra semanal. Para isso, foram construídos cenários com a devida correlação temporal entre os postos horários e dias da semana, o que permite o cálculo dos indicadores de atendimento, dado um número pré-estabelecido de atendentes disponíveis. Por fim, foi implementada uma ferramenta para cálculo dinâmico de volumetria, em que o usuário pode ajustar determinados parâmetros como janela temporal e número de atendentes para obtenção da projeção desejada.

*Keywords:* Customer Service, Average Talk Time, Operational Efficiency, Erlang-C, Kernel, Stochastic Projection, Volumetry, Call Center.

*Palavras-chaves:* Serviço de atendimento, Tempo Médio de Atendimento, Eficiência Operacional, *Erlang-C*, *Kernel*, Projeção Estocástica, Volumetria, *Call Center*

## 1. INTRODUÇÃO

Atualmente, o setor de distribuição de energia passa por significativas mudanças de paradigmas, como a disseminação de recursos energéticos distribuídos, estabelecidas em uma regulação por incentivo que considera um ganho anual de produtividade, o que aumenta a complexidade da rede e reduz o faturamento das distribuidoras, exigindo

uma gestão de processos mais eficiente e esforços, cada vez maiores, para reduzir gastos operacionais. Nesse âmbito, dimensionar corretamente os serviços terceirizados a serem contratados – como o *Call Center* – é essencial para que se evite investimentos ineficientes, e se garanta a viabilidade do negócio de distribuição, bem como a modicidade tarifária.

Nos diferentes contextos em que se inserem serviços regulados, tais como distribuição de energia e telefonia, os atendimentos via *Call Center* representam um direito regulado e

<sup>\*</sup> Este estudo foi financiado e realizado no âmbito do projeto de P&D ANEEL PD-00063-3037/2018.

de cumprimento obrigatório que deve ser garantido ao consumidor. Nesse âmbito, essa estrutura centralizada destinada a receber ligações telefônicas tem caráter primordial, e demanda não apenas de investimentos bem direcionados, mas gestões apropriadas de processos, que assegurem a qualidade e eficiência dos atendimentos. A estruturação, assim como a efetividade desses serviços sofrem alterações contínuas à medida que novas tecnologias são incorporadas nos processos e eles se adequam a oportunidades apresentadas através de análises comportamentais e de procedimentos, associadas às dinâmicas em que os agentes e consumidores se inserem (Aksin et al. (2007)).

O *Call Center* das distribuidoras da CPFL Energia, um serviço regulado pela ANEEL, funciona em regime ininterrupto e, atualmente, possui modelo de gestão terceirizado, sendo um serviço contratado pelas distribuidoras. A contratação do volume de atendimento deve respeitar algumas determinações regulatórias previstas na Resolução da ANEEL nº 414/2010, como o percentual de chamadas atendidas em até 30 segundos em relação ao total de chamadas atendidas (regulado pelo indicador INS, Índice de Nível de Serviço), o percentual de chamadas abandonadas em tempo superior a 30 segundos em relação às chamadas atendidas (regulado pelo Índice de Abandono, IAb) e o percentual de chamadas ocupadas em relação ao total de chamadas oferecidas (monitorado pelo ICO, Índice de Chamadas Ocupadas).

A contratação de chamadas de *Call Center* atualmente é feita pela CPFL Energia com base no método de *Speaking Time*, que leva em conta a quantidade de minutos falados em seus acordos contratuais. No entanto, esse método pode levar a uma superestimação da volumetria contratada, incorrendo em tempo ocioso dos operadores e investimento ineficiente. Essa conjuntura faz com que o cálculo dos minutos de serviços de *Call Center* a serem negociados em contrato requeira análises descritivas minuciosas e projeções assertivas, a fim de proporcionar a redução de custos e garantir a prestação de serviços de qualidade capazes de atender as demandas necessárias ao mercado consumidor.

Dado o contexto de funcionamento e estruturação dos atendimentos telefônicos, o objetivo deste estudo consiste em reduzir os custos operacionais do *Call Center* e mitigar riscos regulatórios. Isso foi proposto através do desenvolvimento de uma solução que aumente a assertividade da volumetria considerada na contratação do *Call Center* – isto é, do dimensionamento correto da volumetria a ser contratada, de modo que o contratado se aproxime ao máximo do realizado e se minimize o tempo de ociosidade dos agentes de atendimento. Resolver este problema é fundamental para garantir a otimização dos recursos empregados em atendimento, de forma a manter a qualidade dos serviços e otimizar os custos.

Para tanto, estruturou-se um banco de dados históricos, a partir do qual um modelo automatizado foi implementado para projeção de demanda através de simulações estocásticas, baseadas em distribuições de *Kernel* e funções de acoplamento. Cenários semanais foram constituídos para intervalos glandulares de 30 minutos, levando em consideração a correlação temporal entre postos horários, dias da semana e implicações ocasionadas pela sazonalidade.

A metodologia de *Erlang-C* foi também aplicada para determinação da demanda ótima estimada de operadores.

Estudos correlatos abordaram o cenário de otimização da contratação dos serviços de *Call Center* fazendo uso de estratégias diversas, tais como aplicações empíricas associadas a séries temporais, em que o método de previsão ARIMA foi aplicado para identificar padrões sazonais (Ferreira (2018)). Analogamente, métodos Regressão Múltipla foram utilizados para implementar modelos preditivos com variáveis binárias capazes de fazerem uma estimativa adequada para previsão de demanda de *Call Center* (Bouzada and Saliby (2009)).

Os métodos de projeção desenvolvidos pela CPFL Energia através desse estudo permitiram o refinamento da estimativa da demanda de minutos contratados, uma vez que abordou, em sua análise, janelas de tempo com maior granularidade e incorporou a interferência de fatores diversos no cenário de atendimentos. Por conseguinte, foi observada uma redução na estimativa dos minutos a serem contratados, além da otimização no cálculo de agentes alocados para efetivação dos atendimentos, garantindo redução de custos e eficiência operacional, fornecendo, ademais, uma ferramenta para a validação dos recursos alocados na empresa contratada para atendimento.

## 2. ANÁLISE DESCRITIVA DA VOLUMETRIA PARA CALL CENTER

### 2.1 Variáveis Explicativas

Dentre as variáveis que influenciam diretamente o fluxo de chamadas e que podem impactar a distribuição das chamadas dentre os dias da semana ou mesmo intra hora (por períodos de 30 minutos), destacam-se: Meteorologia, Absenteísmo, Demandas Comerciais, Retenção de URA (Unidade de Resposta Audível).

*Meteorologia:* Condições climáticas, desde uma perspectiva mais geral de padrões descritos pela sazonalidade anual, que define meses com maiores probabilidades de umidade, até condições ou eventos climatológicos pontuais, têm relação direta com ocorrências na rede elétrica, e portanto, com o número de chamadas recebidas nos serviços de atendimento da distribuidora. É constatado, para períodos chuvosos ou durante temporais, maiores números de interrupções no fornecimento ou ocorrências elétricas diversas registradas, devido a perturbações causadas na rede de distribuição (por exemplo, pela interferência de vegetação), em equipamentos técnicos específicos que constituem o sistema de distribuição ou outros tipos de acidentais. Sendo assim, a questão meteorológica tem grande relevância no que diz respeito ao volume de chamadas tratadas pelos serviços de *Call Center*, uma vez que esses distúrbios levam o consumidor a contatar a distribuidora para reportar esses episódios ou buscar por informações.

*Absenteísmo:* Eventuais ausências pontuais na equipe e redução na capacidade operativa para atender chamadas de *Call Center* podem interferir no potencial de atendimento de chamadas.

*Demandas Comerciais:* No contexto de uma distribuidora, diversos são os serviços comerciais a serem disponi-

bilizados e atendidos ao consumidor. Quando o consumidor necessita de um serviço de caráter comercial, seja ele uma ligação nova, um desligamento, uma solicitação de PID (pedido de indenização de danos elétricos), alguma atuação relacionada ao seu faturamento, leitura, entre outros, ele contata a distribuidora através de seus diversos canais de atendimento. Nesse contexto, o volume de demandas comerciais em um determinado período de tempo está diretamente relacionado com o número de chamadas recebidas pelo *Call Center*.

**Retenção de URA:** Na estrutura de atendimento telefônico da CPFL Energia, toda chamada realizada pelo consumidor é direcionada, automaticamente, para a URA, que consiste em um atendimento eletrônico capaz de responder dúvidas ou fornecer informações sem a intervenção de um atendente. Todavia, caso o consumidor necessite da intervenção humana, ele pode ser direcionado, através do fluxo da URA, para o *Call Center*, no qual será atendido por um profissional preparado para tratar a sua solicitação. Nesse contexto, o termo retenção de URA ( $R_{URA}$ ) – em outras palavras, encerramento pela URA – refere-se à capacidade da URA, por si só, solucionar ocorrências, de forma a não demandar a atuação do *Call Center*. Para avaliação desse fator, são consideradas as variáveis: volume das chamadas oferecidas (COF), ocupadas (CO) e recebidas (CR).

Com o objetivo de aumentar esse indicador, foi implementada para todas as distribuidoras, em fevereiro de 2017, uma nova plataforma de atendimento. Através do incremento na capacidade de recebimento de chamadas simultâneas pela URA (de 240 para 490) essa iniciativa proporcionou um aumento significativo na Retenção de URA. Não obstante, o volume de chamadas recebidas e atendidas pelo *Call Center* se manteve constante, indicando uma independência destes parâmetros quanto a  $R_{URA}$  – cujo incremento permitiu, por outro lado, que mais consumidores fossem atendidos simultaneamente pela URA. Por esta razão,  $R_{URA}$  não é incluída no modelo de projeção da volumetria do *Call Center*. No entanto, as variáveis históricas utilizadas para seu cálculo – em conjunto com outras, tais como o volume de chamada atendidas (CA), abandonadas (CAB), o tempo médio de atendimento (TMA), o número de atendentes e os próprios indicadores regulados (INS, IAb, ICO) – foram consideradas em diferentes etapas dos estudos desenvolvidos nesse trabalho.

Como exposto, são diversos os fatores capazes de interferir na efetividade operacional dos serviços de *Call Center*. É importante também enfatizar que eventos pontuais, tais como interrupções ou reduções operativas em outros canais de atendimento (site, aplicativo, agências físicas) impactam na demanda por serviços de atendimento telefônico.

## 2.2 Análise Estatística do Histórico de Chamadas Atendidas

Foram realizadas análises do histórico de chamadas atendidas para todas as distribuidoras do Grupo CPFL Energia. Para fins de apresentação e análise, neste documento alguns resultados apresentados são referentes exclusivamente à empresa CPFL Paulista.

A contratação de volumetria utilizada anteriormente a esse estudo era feita através de diagnósticos sobre planilhas

históricas consolidadas com base nos fatores: período do ano e dias da semana. Sobre essas variáveis, eram feitos agrupamentos: **(I)** por períodos: 1 - úmido (dezembro, janeiro, fevereiro e março); 2 - seco (abril, maio, junho, julho e agosto); 3 - intermediário (setembro, outubro e novembro); e **(II)** por dias da semana: a - dia de semana (dias úteis: segunda, terça, quarta, quinta e sexta); b - sábado; c - domingo. A influência desses critérios tradicionalmente considerados, podem ser verificados através da evolução temporal no volume de chamadas recebidas, apresentada em Fig.1, para o período de 2013 a 2018. Como pode ser visualizado, há uma sazonalidade entre os períodos seco, úmido e intermediário, com considerável aumento do volume de chamadas no período úmido.

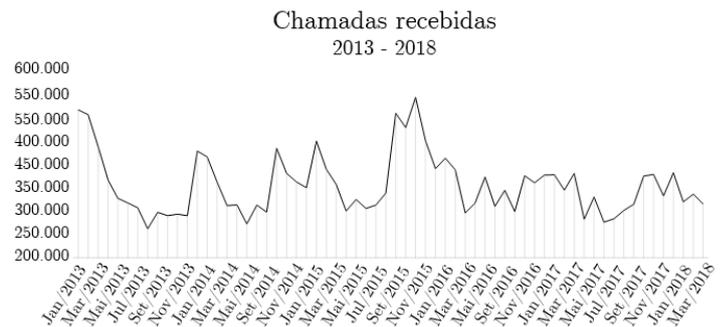


Figura 1. Histórico de chamadas recebidas por mês - CPFL Paulista.

Para a análise do volume de chamadas considerando as variações quanto aos dias da semana, a Fig.2 apresenta o perfil médio semanal de chamadas recebidas do período úmido – maior volume médio e, por isso, destacado. Dentre os dias úteis do período úmido, os maiores volumes de atendimento estão entre segunda e quarta-feira com aproximadamente 17,5% por dia do volume semanal total. Após quarta, o volume apresenta um ligeiro decaimento: quinta-feira (17%) e sexta-feira (15%). Por fim, fins de semana representam apenas 9% e 6,5% do volume de chamadas semanal total, para sábado e domingo, respectivamente. Nesse contexto, ressalta-se que como o perfil e volume de chamadas entre segunda e sexta-feira são semelhantes, o volume de chamadas contratadas em um mês é igual de segunda a sexta-feira, e para sábados e domingos, há especificação de volumes distintos. No agrupamento, em dia útil (dia de semana) do período úmido 71% do volume de atendimento diário ocorre dentro do horário comercial, sendo 32% é pela manhã, com pico de 4% do volume diário às 10h30.

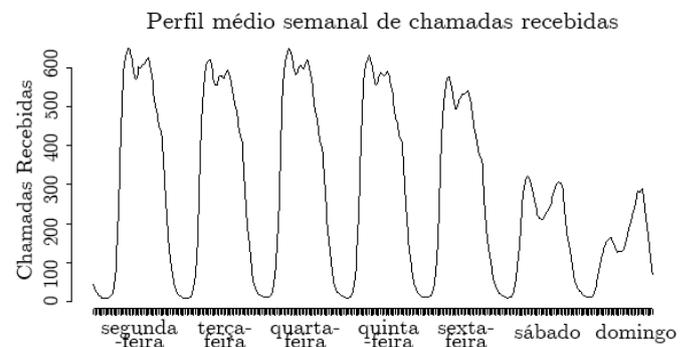


Figura 2. Média semanal de chamadas recebidas para período úmido - CPFL Paulista.

### 3. METODOLOGIA

#### 3.1 Projeção Estocástica do Volume de Chamadas Recebidas

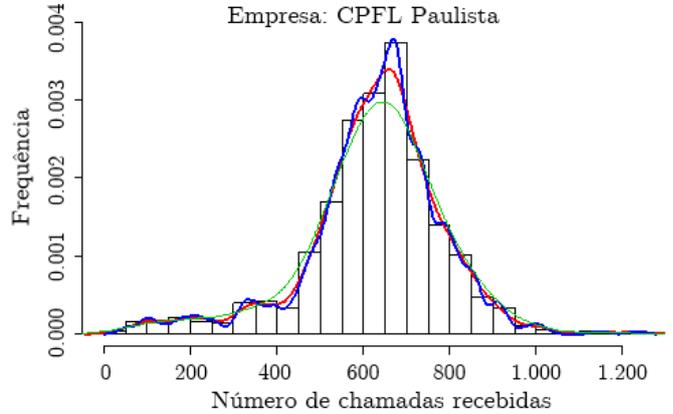
Como demonstrado na análise descritiva, os dados de chamadas recebidas possuem uma grande variação, de forma que modelos baseados em médias, como o utilizado até o momento, não permitem um controle do risco a que a distribuidora está exposta, uma vez que não levam em consideração as possibilidades de ocorrências especificamente para determinados horários, dias, semanas e meses. Desta forma optou-se por uma abordagem estocástica do problema. A estatística descritiva apresentada no tópico anterior fornece uma percepção do funcionamento estocástico do recebimento dos contatos ao longo do tempo.

Um modelo estocástico para um determinado conjunto de eventos aleatórios (Pinsky and Karlin (2010)) deve ser capaz de ajustar, com um grau mensurável de erro, as realizações do espaço amostral (conjunto possível de resultados de uma variável aleatória ou experimento estocástico). No caso desse estudo, o espaço amostral é o número de chamadas no intervalo de tempo de 30 minutos, que tem características distintas (realizações possíveis) para cada um dos 48 intervalos de 30 minutos diários, com distinções por dia da semana e meses de períodos úmido, seco ou intermediário. Existem, no entanto, opções de distribuições de probabilidade não paramétricas (Siegel and Castellan Jr (1975)) – empírica, *Splines*, *Wavelets*, *Kernel*, etc. – que não fazem distinção sobre a forma da distribuição de probabilidade, e se ajustam conforme histórico de realizações da variável aleatória para se aproximar à forma de uma distribuição de probabilidade desconhecida, ou que poderia ser de maior dificuldade de ajuste paramétrico devido à falta de número suficiente de observações históricas. Essa metodologia é flexível, e se ajusta com qualidade, mesmo quando a distribuição observada se afasta consideravelmente das expectativas de qualquer distribuição paramétrica. Por esta razão, tendo em vista que as distribuições observadas para as janelas de 30 minutos não apresentam formas paramétricas de distribuição, este estudo opta pelo modelo estocástico de distribuição *Kernel* (Ma and Zabaras (2011)) com o objetivo de representar as características de volumetria e TMA de um determinado intervalo de tempo. Essa metodologia permite simulações e projeções mais próximas às características reais do intervalo, abrangendo a possibilidades não contidas no histórico, conforme reproduzido na Fig.3.

Estimar as distribuições de probabilidade individuais para cada período de 30 minutos, por dia da semana e período anual é fundamental para se obter as estimativas e projeções granulares nesses intervalos. Não obstante, ao observar os dados históricos nota-se uma correlação temporal importante entre os intervalos de 30 minutos e entre dias de semana. A Fig.4 exibe a correlação temporal entre os períodos intra horários, a cada 30 minutos de uma semana com cada um desses mesmos horários e dias da semana. Quanto mais alta a correlação, maior será a temperatura da cor no gráfico. A primeira diagonal do gráfico exibe a correlação entre as horas de um mesmo dia (por exemplo, as correlações da primeira até a última hora da segunda-feira). Esse comportamento temporal deve ser utilizado

para compor o modelo estocástico de projeção e simulação, de forma a representar dias e semanas com elevações e reduções no volume de chamadas recebidas.

Histograma de chamadas recebidas por dia da semana às 10:30



Chamadas recebidas às 10:30h em dias de semana de meses do período seco, ajustada por distribuições *Kernel* com diferentes bandas de alisamento, linhas: verde (alisamento duplo), vermelha (alisamento padrão) e azul (alisamento reduzido 50%).

Figura 3. Chamadas recebidas às 10:30h em dias de semana de meses do período seco.

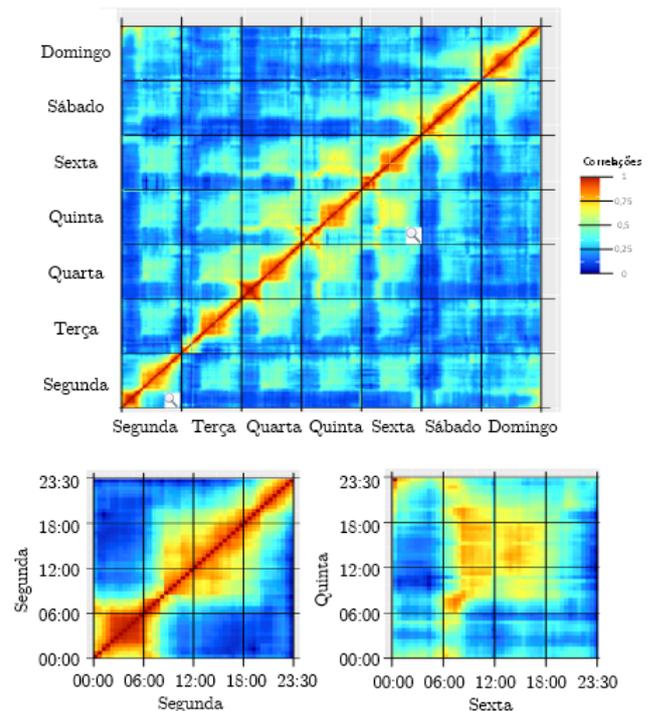


Figura 4. Correlações temporais entre períodos de 30 minutos. Detalhe ampliado: correlações horárias entre as horas de uma segunda-feira e uma sexta-feira.

O modelo estocástico de distribuição *Kernel* para um determinado horário é tratado independentemente dos modelos estocásticos de outros horários, de forma que se deve inserir o comportamento correlacional a essas distribuições, para que as simulações de volumes horários de chamadas

possam ser condizentes com o comportamento temporal observado no histórico. Para tanto, é utilizado o conceito de funções de acoplamento. Com os modelos marginais e a estrutura de correlação temporal, foi obtido o modelo de simulação semanal, que gera cenários em base intra horária (períodos de 30 minutos) para cada um dos 7 dias da semana, com dimensão total de 336 pontos temporais (48 cenários por dia  $\times$  7 dias por semana) correlacionados. Com esse modelo é possível simular uma semana de qualquer um dos períodos anuais (períodos úmido, seco e intermediário).

A geração aleatória de cenários irá produzir inicialmente 2.000 cenários distintos e possíveis de chamadas recebidas para essa semana padrão com base em premissas selecionadas previamente (como período do ano ou mês). Para a simulação de uma operação mensal, será desconsiderada a correlação temporal entre semanas, dando ao modelo a possibilidade de haver semanas de alto ou baixo volume de chamadas dentro de um mesmo mês, ainda respeitando a correlação temporal intra semanas (entre os diferentes dias de uma semana) e intra horários (entre os diferentes horários de um dia).

A partir do conjunto de simulações, foi possível estabelecer diferentes abordagens para o estudo: (i) o uso das simulações para aferição dos riscos de ultrapassagem das métricas de controle regulatórias; (ii) o uso das médias das simulações para projeção do volume médio mensal; (iii) o cálculo semelhante ao realizado atualmente na contratação mensal, porém com metodologia estocástica; (iv) e a projeção através de percentis horários, semanais ou mensais. A projeção por percentis baseia-se na utilização da distribuição de probabilidade e do conjunto de cenários simulados, para verificar se existe algum valor que mantém um percentual definido de realizações possíveis menores ou iguais a esse percentil. Definido, por exemplo, o percentual de interesse como 50%, deseja-se um volume de chamadas para um determinado horário com 50% de chance de não superar o percentil teórico ou simulado escolhido.

### 3.2 Projeção Estocástica para TMA

Assim como realizado para o volume de chamadas (seção 3.1), o TMA também possui seu grau de variabilidade, que pode ser investigado, simulado e projetado para cada período horário, dia da semana e período mensal. A partir da análise descritiva do TMA, observou-se uma aproximação das distribuições à distribuição Normal, porém com desvios da normalidade em alguns casos. Assim, também é útil a flexibilidade da distribuição *Kernel* para o TMA, já que irá ajustar os casos de desvio de normalidade. Com os modelos marginais *Kernel* para o TMA e a estrutura de correlação temporal, foi obtido o modelo de simulação semanal, que gera cenários em base intra horária (períodos de 30 minutos) com dimensão total de 336 pontos temporais correlacionados. Podendo as simulações, assim como no caso do volume de chamadas, serem utilizadas para a projeção por volume médio mensal, percentil horário, semanal ou mensal.

### 3.3 Modelo Automatizado para Projeção Semanal de 48 Postos Diários e Número de Atendentes

Utilizando a função de distribuição *Kernel* e função de acoplamentos, foi desenvolvido, assim como descrito em

3.1, um modelo automatizado para a projeção semanal de 7 dias da semana e 48 postos (ou janelas/intervalos) diários, totalizando 336 postos simulados para uma semana padrão, com duração de 30 minutos cada um. São produzidos, a partir de simulações estocásticas, 2.000 cenários que possuem a devida correlação temporal entre postos horários e dias da semana. Dessa forma, com o objetivo de projetar de maneira automática o volume de chamadas para um determinado cenário semanal com seus 48 postos diários, é feita a interseção de um conjunto de dados históricos adequados: visando simular uma projeção semanal para o mês de outubro de 2018, seleciona-se como *input* as observações de outubro de 2014 a 2017, que reproduzirão 2.000 cenários de uma semana padrão para esse mês e, através das médias dos postos (ou janelas) simulados, obterá a projeção do volume de chamadas para essa semana.

Esses conjuntos de 2.000 cenários consolidados permitem também a estimativa do número teórico de operadores, através da implementação de um algoritmo baseado no modelo de filas generalizadas que utiliza a técnica *Erlang-C* (Angus (2001)). Esses cálculos permitem estabelecer, de forma automatizada, o número de atendentes necessários para suprir um determinado volume de chamadas médio/contratado.

### 3.4 Simulação de Atendimento em Tempo Real e Cálculo de INS Resultante

Para permitir uma estimativa dos indicadores de qualidade do serviço de atendimento mais confiável, é preciso simular a operação do *Call Center* em um dia qualquer. Considerando que o *Call Center* possui um número variável de atendentes ao longo de um dia, e quantidades incertas de chamadas, foi construído um sistema de simulação em tempo real, com o objetivo de simular, segundo a segundo, o atendimento de chamadas dado uma expectativa de número de atendentes em cada posto intra horário (48 postos de 30 minutos em um dia, totalizando 86.400 segundos). Esse sistema de simulação recebe como premissas:

- (1) Um cenário simulado de chamadas recebidas, para cada horário do dia sendo simulado;
- (2) Um cenário simulado do TMA das chamadas para cada horário do dia sendo simulado;
- (3) Distribuição de probabilidade de TMA; e
- (4) Número de atendentes esperado para atendimento do volume de chamadas.

Com essas informações, a cada intervalo de 30 minutos, distribui-se as chamadas simuladas para o intervalo, de acordo com determinada regra de distribuição, ou uniformemente ao longo dos segundos que o compõe, sendo atribuído, a cada chamada, um tempo de atendimento simulado, cuja média é o TMA estimado para aquele horário. Em seguida, o algoritmo aloca, a cada atendente teórico, uma chamada, registrando o segundo do início e o segundo de encerramento da chamada. Caso não haja mais atendentes para atender uma chamada simulada em um determinado segundo, quando todos estiverem ocupados, o algoritmo registra segundo a segundo o tempo de espera, até que aquela chamada seja atendida. O algoritmo é versátil para ser utilizado em qualquer dia de qualquer semana ou mês, bastando informar o número de atendentes e o cenário de chamadas a ser calculado.

### 3.5 Validação e Calibragem dos Modelos

Com os modelos de simulação de volume, simulação da operação e *Erlang-C* (número de atendentes) construídos, foi feita a validação e a calibragem do modelo, juntamente com a simulação da lista ordenada do portfólio e empacotamento do *input* dos parâmetros para dimensionamento do *Call Center*. A calibragem foi realizada para que os resultados obtidos de INS e IAb estivessem próximos dos realizados ao longo dos meses de janeiro de 2018 a julho de 2018. Os modelos foram todos desenvolvidos em R e posteriormente os algoritmos foram empacotados para executar por meio de uma Macro em Excel, de forma a atender as demandas requeridas para implementação na área de negócio. Essa abordagem reduziu o tempo necessário para que o modelo fosse colocado em produção, assim como aquele demandado para treinamento do operador, o qual pôde continuar utilizando uma ferramenta com a qual já tem familiaridade.

## 4. RESULTADOS

Como resultado da criação, desenvolvimento e calibragem foi obtido o modelo estocástico automatizado, que considera as correlações temporais e metodologias estatísticas. O modelo utiliza como base para projeção o comportamento histórico dentro do período designado nos campos “Ano Início Histórico” e “Ano Fim Histórico”, utilizando somente as informações constantes no “Mês Referência”, ou seja, projetando a volumetria para o mês 10/2018 (conforme Fig.5), cujo período de base é de 2015 até 2017, o modelo utilizará os meses 10/2015, 10/2016 e 10/2017 para projeção da volumetria de 10/2018.

Neste modelo, a atualização é automática, onde basta selecionar o mês e ano de referência, o período histórico como base da projeção e, se necessário, a calibragem dos horários fora ponta (entre as 18h – 7h) e horários de ponta (entre as 7h – 18h), e acionar o botão “Rodar Volumetria”, iniciando a geração da volumetria e o número de atendentes estimados, individualizados por empresa e separados por intervalos de 30 minutos. A Fig.5 ilustra a ferramenta implementada para as projeções, que permite que o usuário possa ajustar as premissas a serem consideradas no cálculo, tais como o período considerado e o número fixo de atendentes disponíveis.

CPFL Energia - Sistema de Simulação de Atendentes e Volumetria					
<b>DADOS BÁSICOS:</b>					
Mês Referência	10	Ano Início Histórico	2015		
Ano Referência	2018	Ano Fim Histórico	2017		
Arquivo Saída	15-17_Simulação_201810_Q5SQ60_95_Atend_Fer.xlsx				
Roda INS esperado	Não				
Usa Atendentes Fixo	Não				
Soma RGE+RGS	Não				
<b>Rodar Volumetria</b>					
<b>DADOS Calibragem:</b>					
Alisamento Kernell	1,5	Dim. Posto Horário	30	S1	1
Quantil TMA	0,8	Sec. Atend. Mínimo	20	S2	1,025
Quantil CA Fora Ponta	0,55	Atend Min %	0,95	S3	0,973
Quantil CA Ponta	0,6	Seed Value	999	S4	0,969
				S5	0,969

Figura 5. Tela para atualização do modelo proposto.

Com o novo modelo foi possível reduzir a granularidade das chamadas, projetando, individualmente, cada dia do

mês. Com a granularidade reduzida para cada dia do mês, o novo modelo projeta, para cada empresa, as chamadas em intervalos de 30 minutos com maior precisão do que os resultados obtidos para os cálculos realizados anteriormente na companhia. Além da estimativa da volumetria de chamadas, algoritmos automatizados permitiram projetar de maneira otimizada o número de operadores necessários para suprir demandas específicas.

Na Fig.6 é apresentado o percentual do desvio total entre os volumes calculados para atendimento do modelo desenvolvido e o volume contratado atual, considerando a quantidade de volume acumulado no período de Março a Agosto de 2018. Na empresa CPFL Paulista, por exemplo, o desvio percentual do modelo foi de 2% acima e do volume contratado foi de 2% abaixo, ou seja, a contratação foi abaixo do esperado para realização dos atendimentos. A empresa que teve o maior desvio percentual foi a RGE, em que o modelo apresentou um desvio de 22% e o do modelo de contratação atual foi de 38% acima, assim, mesmo esse sendo o maior desvio apresentado pelo modelo, sua performance é melhor, representando uma diferença entre eles de 16 pontos percentuais. Em todas as observações, mesmo sem utilizar parâmetros individualizados para as empresas, o modelo apresenta maior precisão do que a metodologia de contratação atual em todas as empresas apontadas.

Comparativo do percentual do desvio total entre o resultado do modelo e o contratado

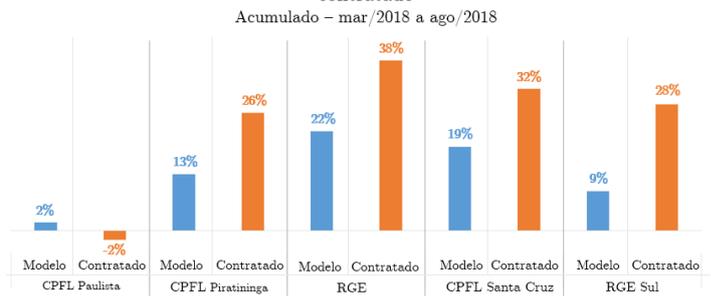


Figura 6. Comparativo do percentual do desvio total entre o modelo desenvolvido e o modelo atual de contratação de volumetria do *Call Center*.

## 5. CONCLUSÃO

Com o intuito de dimensionar corretamente os serviços terceirizados a serem contratados, neste caso específico o serviço de *Call Center*, é essencial que se evite investimentos ineficientes, e se garanta a viabilidade do negócio de distribuição de energia elétrica, como com a modicidade tarifária. Para isso, o Centro Analítico – departamento da companhia financiado por um projeto P&D da ANEEL e destinado a criar valor a partir da integração e análises de dados – propôs uma projeção otimizada para cálculo de volumetria de serviços de *Call Center*.

No modelo original eram desconsideradas as variáveis que podem influenciar na variação da volumetria da contratação de atendimento, tendo por consequência uma subcontratação ou sobrecontratação e que poderia impactar nos indicadores regulados de maneira negativa. O cálculo da volumetria era realizada por dias da semana (dias

úteis, sábados e domingos) com 18 meses de antecedência, junto ao orçamento da área responsável pela contratação, considerando a média dos 6 maiores picos dos últimos 18 meses, a realização histórica dos últimos 4 anos, ou a média da volumetria para o período de referência.

Visto isso, foram desenvolvidos os modelos de projeção estocástica do volume de chamadas recebidas e do TMA, modelo de simulação de atendimento em tempo real e cálculo de INS resultante e, por fim, o modelo de projeção semanal de 48 postos diários. Nos modelos de projeção estocástica foi projetado um mês especificado, por intervalo de 30 minutos, considerando um nível de risco definido. Já no modelo de simulação de INS, é um algoritmo versátil que pode ser utilizado em qualquer dia de qualquer semana ou mês, bastando informar o número de atendentes o cenário de chamadas a ser calculado. Finalmente, no modelo de projeção semanal de 48 postos diários foi implementado um algoritmo de forma a alocar a quantidade de atendentes necessários para suprir um determinado volume de chamadas médio/contratado.

Com isso o novo modelo utiliza a projeção estocástica que considera as correlações temporais e metodologias estatísticas. A sua atualização é feita de maneira automática, tem-se em relação a granularidade uma projeção das chamadas em intervalos individuais de 30 minutos. A parametrização do risco é feita com base no percentil de probabilidade de realização para horário de ponta e fora de ponta. Por fim, no modelo original não era feita a estimativa do número teórico de atendentes, o que foi incorporado no novo modelo, em que a estimativa é feita com base em modelo de filas generalizadas.

## REFERÊNCIAS

- Aksin, Z., Armony, M., and Mehrotra, V. (2007). The modern call center: A multi-disciplinary perspective on operations management research. *Production and operations management*, 16(6), 665–688.
- Angus, I. (2001). An introduction to erlang b and erlang c. *Telemanagement*, 187, 6–8.
- Bouzada, M.A.C. and Saliby, E. (2009). Prevendo a demanda de ligações em um call center por meio de um modelo de regressão múltipla. *Gestão & Produção*, 16(3), 382–397.
- Ferreira, A.C.d.S. (2018). *Previsão de séries temporais com períodos sazonais complexos: uma aplicação empírica*. Ph.D. thesis, Instituto Superior de Economia e Gestão.
- Ma, X. and Zabaras, N. (2011). Kernel principal component analysis for stochastic input model generation. *Journal of Computational Physics*, 230(19), 7311–7331.
- Pinsky, M. and Karlin, S. (2010). *An introduction to stochastic modeling*. Academic press.
- Siegel, S. and Castellan Jr, N.J. (1975). *Estatística não-paramétrica para ciências do comportamento*. Artmed Editora.