

Análise de Cadastros de Concessionárias

- Metodologia -

SÃO PAULO, 2000

CONVÊNIO ANEEL / FUSP



Convênio ANEEL/FUSP: Estudos e Pesquisas Sobre Regulação e
Fiscalização do Setor Elétrico Brasileiro.
Projeto: Metodologia de Análise de Cadastros de Concessionárias

Índice

Parte I – Estrutura dos Cadastros

Parte II – Algoritmos de Controle

Parte III – Caracterização dos Dados para Aplicação dos Algoritmos

Parte IV – Séries Temporais



Convênio ANEEL/FUSP: Estudos e Pesquisas Sobre Regulação e
Fiscalização do Setor Elétrico Brasileiro.
Projeto: Metodologia de Análise de Cadastros de Concessionárias

Estrutura dos Cadastros



Convênio ANEEL/FUSP: Estudos e Pesquisas Sobre Regulação e
Fiscalização do Setor Elétrico Brasileiro.
Projeto: Metodologia de Análise de Cadastros de Concessionárias

Índice

Introdução	3
1. Origem do cadastro	4
2. Caracterização dos Cadastros Recebidos	7
2.1. Estrutura	7
2.2. Comentários e recomendações	8



Convênio ANEEL/FUSP: Estudos e Pesquisas Sobre Regulação e
Fiscalização do Setor Elétrico Brasileiro.
Projeto: Metodologia de Análise de Cadastros de Concessionárias

Introdução

Os cadastros comerciais dos concessionários, registrando os contratos de suprimento e os consumos, mês a mês, dos diferentes consumidores, representam um precioso instrumento de fiscalização das condições operativas dos concessionários. Utilizando-se de algoritmos de varredura automática dos campos armazenados eletronicamente, é possível verificar o cumprimento de uma série de obrigações dos concessionários, identificar práticas não corretas e ajudar na identificação de situações anômalas, que podem colocar em risco a qualidade da alimentação elétrica.

Apesar do grande potencial oferecido para fiscalização, atualmente os concessionários obtêm dos cadastros somente relatórios que utilizam os dados de forma agregada, gerando tabelas de resumo da situação do mercado ou tabelas com os resumos de distribuição dos consumidores por faixa de consumo.

O uso do cadastro para controle e fiscalização exige porém, uma compreensão mais aprofundada dos campos contidos no cadastro e a verificação do preenchimento adequado destes para contemplar as exigências do fiscalizador.

Este manual contém uma análise da estrutura conceitual dos cadastros e comentários sobre alguns cadastros específicos conhecidos. Considerando que o detalhamento das informações contidas no cadastro é em grande parte dependente da legislação, incluiu-se um comentário sobre os instrumentos legais que formulam estas exigências, especificamente a portaria 466 (Contribuições para a Revisão da Portaria 466 do DNAEE, de 12 de novembro de 1997 - manual I, volume II) .

1. Origem do cadastro

O sistema elétrico é fundamentalmente um sistema de distribuição em grande escala sobre uma base geográfica. Não sendo a energia elétrica um produto armazenável, o sistema deve necessariamente levar em conta duas variáveis: a variável geográfica, da distribuição territorial dos consumidores e a variável temporal, representada pela demanda específica de cada consumidor, que varia ciclicamente durante o dia, a semana e o ano, além disso, pode apresentar uma variação tendencial, em geral crescente, ao longo dos anos.

Para documentar as variáveis geográfica e temporal, essenciais para o planejamento das linhas e para o faturamento, os cadastros nasceram como livros contábeis, nos quais eram agrupados os consumidores. Cada livro ou razão representava uma área geográfica que agrupava um número variável de rotas ou roteiros. O leiturista possuía um livrete, no qual anotava o consumo, e cada livrete agrupava o número de consumidores que o leiturista conseguia ler em um dia de trabalho. Assim, o registro mensal do consumo de energia de cada consumidor, os livros, os roteiros e os livretes, representavam, a grosso modo, a distribuição do consumo em base geográfica e sua evolução no tempo, identificando alimentadores, distribuição, transformadores, etc.

Na realidade, com o passar do tempo, a base geográfica do cadastro elétrico acabou se afastando da base geográfica da distribuição e assumiu o caráter de simples livro de faturamento (separação entre os alimentadores e a distribuição gerando, desta forma, uma separação entre as necessidades do faturamento e as técnicas), mantendo porém, algumas características que remontam de seu papel enquanto livro contábil. As principais razões desta evolução são:

a. As estruturas urbanas crescem frequentemente sem levar em conta a distribuição dos serviços essenciais, tornando instável a base geográfica sobre a qual é montada o cadastro. A eletricidade é um dos primeiros serviços públicos fornecidos em bairros novos e isso faz com que o contrato possa ser estipulado antes do arruamento e portanto da numeração das casas. A consequência disto, em algumas áreas de distribuição, é que pode-se encontrar até três numerações e três nomes de ruas diferentes para o mesmo endereço físico: endereço dos serviços elétricos, endereços dos serviços da água e endereço oficial da prefeitura. Esta situação, para garantir a leitura mensal regular dos consumos, obriga o concessionário a utilizar o roteiro como verdadeiro guia geográfico, no qual o código de leitura é um número

sequencial, que representa a ordem que o leiturista segue para leitura. Às vezes o roteiro é até representado com uma seqüência de códigos colocados sobre um mapa, que representa o percurso físico do leiturista no processo de controle mensal.

b. As necessidades da leitura são bastante variáveis, dependendo da densidade de consumidores e da facilidade de acesso à leitura. Em regiões com alta densidade de consumidores, por exemplo, em edifícios de apartamentos, o leiturista pode superar quinhentas leituras diárias, mas em regiões de baixa densidade o roteiro pode cobrir um número reduzido da ordem de menos de cem consumidores por dia. A situação da leitura é ainda complicada pelo tipo de padrão de entrada. Em regiões onde a eletrificação é antiga, o padrão de entrada é, muitas vezes com o medidor na parte interna da residência, exigindo a presença do consumidor para que a leitura possa ser realizada. Nestas condições, além da baixa densidade de carga, o leiturista se depara com a necessidade, para realizar a leitura, de ter acesso ao domicílio, portanto de atrasar o processo de leitura e, em alguns casos a ter a que realizar uma estimativa, quando ele encontrar o domicílio fechado.

Para garantir a estabilidade do faturamento e que ele espelhe realmente a energia consumida pelo consumidor, a portaria 466 prescreve a leitura mensal dos medidores, com um intervalo de aproximadamente 30 dias entre uma leitura e outra. Como reflexo da época na qual os medidores estavam ainda na parte interna do domicílio, a portaria prescreve também, que o concessionário divulgue, no início do ano, um calendário de leitura, pelo qual o consumidor é informado dos dias nos quais a leitura será realizada. Alterações no calendário de leitura devem ser comunicadas com antecedência ao consumidor. (portaria 466, art. 36). É importante notar como o calendário de leitura representa, na realidade, um anacronismo em relação à obrigatoriedade da instalações do medidor externo, já que, quando ele for colocado em posição acessível ao leiturista, o consumidor não precisa estar presente no ato da leitura.

Associados ao mecanismo de livros, razão e libreto, o concessionário agregava as características contratuais do consumidor (tipo de ligação, demanda contratada, e similares), transformando-se assim o sistema em um verdadeiro espelho da distribuição, em seus aspectos comerciais, mas também técnicos.

Com a evolução de sistemas computacionais, o setor elétrico adotou de maneira crescente, uma integração das informações da empresa, unindo, direta ou

indiretamente, todas as informações obrigatórias para o cumprimento da legislação: as informações estritamente comerciais (como identificação do consumidor, local de cobrança, domicílio bancário e similares), e as informações estritamente técnicas (como alimentador, transformador, poste). Dependendo do sistema computacional utilizado e do tamanho da empresa em relação ao número de consumidores, o cadastro assumiu complexidade e configurações distintas. Hoje, no país, existe uma grande dispersão na maneira de caracterização e de organização dos consumidores, que pode basicamente ser dividida em três grupos distintos:

a. Cadastro tradicional.

Existem ainda, casos de cadastros integralmente manuais, com livros, razões e libretos físicas, nos quais os dados são transcritos manualmente, e existe uma estrutura técnica totalmente separada da estrutura de faturamento. São as últimas heranças do passado, que estão prestes a ser abandonadas.

b. Cadastros eletrônicos.

Não existe a integração do cadastro do faturamento com o cadastro técnico. Estes cadastros representam simplesmente o registro dos contratos, com suas características comerciais de identificação do consumidor, tarifa, local de cobrança e a seqüência histórica de consumos e demandas (quando aplicável). Na maioria dos cadastros eletrônicos existe hoje um início de integração com o cadastro técnico.

c. Cadastros georeferenciados.

Os cadastros georeferenciados são aqueles onde uma base de dados cadastrais é ligada a uma base de dados geográfica. A base de dados cadastrais nada mais é que o cadastro de consumidores com um campo para especificar o local do consumidor. Esse campo é relacionado através de um Sistema de Informações Geográficas (SIG) à localidade, em planta, do consumidor. O avanço tecnológico em capacidade de processamento de computadores, em bancos de dados relacionais e em integrações dados-sistema gráfico, permitem ao fiscalizador cruzar informações cadastrais, socioeconômicas e geográficas para detectar os vieses de consumo e localizar numericamente e graficamente os consumidores envolvidos. O nível de complexidade e de integração desses SIGs varia de concessionário para

concessionário pois tecnologia e *know-how* se difundem mais lentamente nos setores operacionais de grandes empresas.

2. Caracterização dos Cadastros Recebidos

2.1. Estrutura

O cadastro do concessionário ELEKTRO já possui elementos de integração entre os elementos técnicos e comerciais que se caracterizam pelo registro dos transformadores e seu controle. A empresa, antes da privatização, utilizava-se o *software* Prodadis para trimestralmente efetuar o cálculo do carregamento provável do transformador.

O cadastro do concessionário BOA VISTA ENERGIA está estruturado para fazer a integração de base georeferenciada, mas ainda não foi efetuada por falta da base geográfica. Está sendo realizado um novo mapeamento da cidade para criar uma cartografia estável que permita a integração.

O cadastro utilizado para os processamentos é um cadastro completo (dadas as limitações do extrator), e representa a população da localidade escolhida (Rio Claro e Boa Vista). Já que a avaliação do cadastro total de uma concessionária não é viável de ser efetuada em microcomputadores, a escolha de um cadastro completo de uma localidade (população) é recomendada porque, enquanto algumas variáveis poderiam ser representadas por uma mostra (por exemplo o consumo médio por consumidor, ou a tendência do consumo), outras variáveis não são passíveis de amostragem. Por exemplo, a regularidade da leitura de uma rota exige o cadastro completo daquela rota (de outra forma os erros de amostragem introduzidos invalidariam o resultado). No tratamento dos dados, cada variável é tratada em relação à população, e as características específicas da população são avaliadas no subconjunto de consumidores que possuem aquela característica, por exemplo, categoria tarifária, rota, classe de atividade.

Os dois concessionários dos quais foi possível obter o cadastro, utilizaram, para fornecer os dados, um extrator, que, dentro da massa de dados disponíveis dentro da concessionária, extraiu unicamente os especificados¹, que correspondem aos requisitos da portaria 466. Os demais dados podem ser considerados pelo concessionário como confidenciais, portanto não existe obrigatoriedade em fornecê-los ao regulador.

Foi possível ainda, obter algumas informações específicas que estão sintetizadas no quadro a seguir.

característica	ELEKTRO	Boa Vista Energia
classe tarifária	código indicando classificação socioeconômica, tipo de ligação e tarifa	código indicando classificação socioeconômica, tipo de ligação e tarifa
rota/roteiro: indica sequência a ser seguida pela equipe de leitura	possui rotas específicas para a leitura de consumidores classificados como poder público, consumidores faturados com tarifa horo-sazonal, consumidores faturados com tarifa do grupo A e consumidores com medição de ativo/reactivo e faturamento e baixa tensão. A rota de leitura dos consumidores industriais em alta tensão é efetuada por equipe da empresa. As demais rotas de leitura são efetuadas por equipe terceirizada.	possui rotas em base geográfica não fazendo distinção por categoria de consumidores.
código de atividade	o setor residencial não possui código numérico de classificação, portanto, não há distinção no cadastro entre um único domicílio, um condomínio, etc. Os demais setores são classificados de acordo com o código C.I.U.U.	o setor residencial não possui código numérico de classificação, portanto, não há distinção no cadastro entre um único domicílio, um condomínio, etc. Os demais setores são classificados numericamente, mas com terminologia diferente do código C.I.U.U.

¹ Os dados recebidos, conforme documentado na correspondência anexa, não eram de outra parte completos, faltando a potência demandada por contrato e a demanda registrada, para os contratos com tarifa binomial, e aparecendo vazios os campos de potência instalada, prescritos na baixa tensão.

2.2. Comentários e recomendações

De acordo com as descrições anteriores, para consulta ou emissão de relatórios padronizados, é necessário pensar em dois caminhos separados:

- a. Análise separada do cadastro do concessionário, que poderá ser realizada unicamente pedindo ao concessionário o cadastro inteiro (nos campos que são sujeitos à fiscalização, já que sobre os outros existem restrições jurídicas de confidencialidade), obtido com um extrator. Sobre este cadastro reduzido, ou uma amostra do cadastro, o fiscalizador pode operar, analisando os dados. Este processo pode ser muito demorado, especialmente em concessionários com grande número de consumidores.
- b. Pedido de relatórios específicos, na base dos algoritmos indicados no manual de algoritmos, que o concessionário pode emitir no momento da varredura do cadastro, desenvolvendo para esta finalidade, programas específicos. O período de abrangência do relatório pode ser anual (análise da tendência e da sazonalidade) ou trimestral (ruído branco, compatibilidade contratual).

Nos dois casos, do ponto de vista jurídico, o concessionário tem a obrigação de emitir relatórios ou de fornecer os dados unicamente dos campos do cadastro pelos quais possui responsabilidade contratual de manter arquivados, basicamente os dados prescritos pela portaria 466, portanto é sobre estes dados que será realizada a análise.

Na prática, porém, o concessionário dispõe, em seus arquivos, de informações muito mais amplas, que poderiam ser muito valiosas para fiscalização, mas que não podem ser exigidas. Por exemplo, a maioria das concessionárias possui rotinas específicas de controle da qualidade da leitura e do faturamento, portanto o algoritmo do ruído branco, proposto, representa na realidade um duplicado dos processos já utilizado pelos concessionários, que tem a finalidade unicamente de assinalar de forma global as situações irregulares.

Por outro lado os concessionários possuem testes de consistência e rotinas de análise bastante avançadas, muitas das quais não são documentadas. Como acontece com frequência nos sistemas computacionais, eles vêm crescendo, dependendo das necessidades dos operadores, e as rotinas efetivamente utilizadas não são incorporadas no manual do sistema, quando este de fato existe. Esta situação gera uma grande dificuldade no conhecimento da filosofia do cadastro, inclusive porque cada setor da empresa, dependendo da sua especialização, tem acesso e

documentação unicamente sobre as rotinas por ele especificamente utilizadas.

Como analisado mais em detalhe no item 2 do manual "Contribuições para a Revisão da Portaria Nº 466 do DNAEE, de 12 de novembro de 1997", a portaria prescreve a obrigatoriedade de um campo, no cadastro, indicando a "carga instalada" (art. 19 parágrafo VII). Esta prescrição não é atendida, já que o consumidor, apesar da obrigatoriedade da notificação de qualquer alteração da carga instalada, de fato não a comunica e os próprios concessionários não incentivam esta praxe, que seria inviável em concessionários com milhões de consumidores.

O campo, em resumo, não é utilizado e o fiscalizador perde qualquer capacidade de controle sobre a adequação do contrato de suprimento utilizado pelo consumidor em relação ao seu consumo ou à sua demanda. Atualmente, com a evolução rápida do mercado, o crescimento da demanda dos consumidores pode levar à condições críticas. Condições estas, como as registradas em algumas concessionárias, provocadas pelo aumento sensível do consumo, que leva a colocar em crise transformadores e alimentadores. Paradoxalmente, estas situações críticas podem ser contestadas pelo concessionário, já que o aumento do consumo se verifica sem que os consumidores notifiquem os aumentos de carga, não caracterizando, portanto, uma responsabilidade do concessionário nas condições normativas atuais.

Uma maneira de contornar o problema seria seguir um outro critério de verificação do problema. Sabendo-se que grande parte das concessionárias está orientada a cadastros em base geográfica, seria possível retirar a obrigatoriedade do registro da carga instalada, criando por outro lado, a obrigatoriedade, por parte do concessionário, em fornecer ao fiscalizador, o consumo acumulado de todos os consumidores, por posto de transformação e por potência instalada de transformação. Com um simples programa de análise, similar àquele proposto para a análise de compatibilidade do contrato, seria possível, para o fiscalizador, verificar concretamente as condições médias de carregamento e, portanto, prever possíveis situações anômalas.

É claro que os concessionários poderiam ter restrições a esta ampliação da fiscalização, porém ocorre lembrar que o fiscalizador se encontra em uma situação de força, já que a atual normativa não é atendida.

Concretamente a compatibilidade do atendimento elétrico dos consumidores em baixa tensão poderia portanto ser realizada com dois algoritmos relativamente simples:

- Compatibilidade contratual, no qual o consumo mensal do consumidor é contrastado com a características do contrato (portanto da potência máxima disponível);

- Compatibilidade do carregamento do transformador, no qual a somatória total dos consumos de todos os consumidores é comparada com a capacidade instalada do transformador.

Algoritmos de Controle



Convênio ANEEL/FUSP: Estudos e Pesquisas Sobre Regulação e
Fiscalização do Setor Elétrico Brasileiro.
Projeto: Metodologia de Análise de Cadastros de Concessionárias

Índice

Introdução	3
1. Glossário	5
2. Algoritmos	6
2.1 Séries temporais	6
Algoritmo 1 - Médias Móveis	6
Algoritmo 2 - Índice de Sazonalidade	9
Algoritmo 3 - Método paramétrico	10
Algoritmo 4 - Ruído Branco	10
Algoritmo 5 - Taxas de Crescimento	11
Algoritmo 6 - Consumo por Consumidor	12
2.2 Cadastro completo dos consumidores (12 meses)	12
2.2.1 Algoritmos para tratamento primário dos dados	12
Algoritmo 7 - Filtragem	12
Algoritmo 8 - Média e desvio padrão do consumo mensal (j)	13
Algoritmo 9 - Média e desvio do consumidor	15
2.2.2 Desconformidades - teste dos outliers	15
Compatibilidade Contratual	15
Algoritmo 10 - Moda	16
Algoritmo 11 - Distribuição de Pareto	16
Algoritmo 12 - Verificação do Contrato no setor Industrial	18
Regularidade do Faturamento	19
Algoritmo 13 - Consumo Mensal Igual a Zero	19
Algoritmo 14 - Índice Consumo Máximo/Consumo Médio	19
2.2.3 Distribuição do consumo de energia	20
Algoritmo 15 - Curva de Lorenz e Índice de Gini	20
Observações	22



Convênio ANEEL/FUSP: Estudos e Pesquisas Sobre Regulação e
Fiscalização do Setor Elétrico Brasileiro.
Projeto: Metodologia de Análise de Cadastros de Concessionárias

Introdução

O presente manual contém a descrição formal dos algoritmos utilizados para análise dos cadastros de concessionárias, no âmbito do projeto "Metodologias de Análise do Cadastro de Concessionárias". No manual III, de análise dos resultados, são apresentados exemplos práticos de aplicação destes algoritmos para os dados de uma concessionária, neste caso a Elektro e mais especificamente para o cadastro da cidade de Rio Claro.

Esta parte especifica os algoritmos (rotinas matemáticas) a serem aplicados aos dados dos cadastros fornecidos pelos concessionários, descrevendo dados de entrada, dados de saída e procedimentos necessários para cada um deles.

Os dados da concessionária são recebidos em um arquivo tipo texto, obtido com um extrator que grava unicamente os campos pertinentes, no qual cada linha corresponde a um consumidor, registrando seu consumo de energia mensal e suas características adicionais: nome, código do consumidor, classe tarifária, bairro, código de atividade, rota, etc

Do ponto de vista computacional, o cadastro é transformado em uma matriz, na qual cada linha corresponde a um registro e cada coluna a um campo. Cada linha, ou registro, identifica um consumidor, identificado pelo código, com as informações complementares contidas nos diferentes campos.

Por exemplo:

X_{ij} - Consumo médio do mês j para o consumidor i , em KWh.

Esta representação é feita porque os dados são apresentados na forma de uma matriz, como a apresentada abaixo.

consumidor	mês 1	mês 2	mês j
1	X_{11}	X_{12}	X_{1j}
2	X_{21}	X_{22}	X_{2j}
...
i	X_{i1}	X_{i2}	X_{ij}

O cadastro utilizado para os processamentos é um cadastro completo (dadas as limitações do extrator), e representa a população da localidade escolhida (Rio Claro e Boa Vista). A escolha de um cadastro completo, ou de uma população, é devida ao fato de que, enquanto algumas variáveis poderiam ser representadas por uma mostra

(por exemplo o consumo médio por consumidor, ou a tendência do consumo), outras variáveis não são passíveis de amostragem. Por exemplo, a regularidade da leitura de uma rota, exige o cadastro completo daquela rota (de outra forma os erros de amostragem introduzido invalidariam o resultado). No tratamento dos dados, cada variável é tratada em relação à população cujas características específicas são avaliadas no subconjunto de consumidores agrupados por parâmetros predeterminados, por exemplo, categoria tarifária, rota, classe de atividade.

1. Glossário

Este glossário tem como objetivo padronizar os termos específicos a serem utilizados neste trabalho.

algoritmo - processo de cálculo de um grupo de dados onde são definidas regras formais para se obter o resultado desejado.

amostra - conjunto de dados extraídos de uma população mediante determinados critérios técnicos que garantem que as conclusões extraídas desta massa de dados possam ser generalizadas para toda a população²

classe tarifária - conjunto de consumidores que possuem o mesmo comportamento de consumo, padrão socioeconômico e padrões de conexão técnica semelhantes (nível de tensão, tipo de contrato de fornecimento de energia).

população - o cadastro completo da localidade/concessionário contendo os registros de consumo de energia, contrato de fornecimento de energia, categoria tarifária, localidade, rota de leitura por um período de doze meses.

séries de consumo - resumo de dados de consumo do concessionário por período preferencialmente de cinco anos ou mais, agregado por nível de tensão ou categoria tarifária.

² Muñoz, M.I., Estadística, 1992, Alhambra Longman Editor

2. Algoritmos

2.1 Séries temporais

Para análise das séries temporais foram adotados dois procedimentos, o método não paramétrico das médias móveis e o método paramétrico com o ajuste a uma função contendo uma componente de incremento e uma componente periódica. O teste das séries pelo método não paramétrico X11 não forneceu resultado significativo que justificasse sua adoção. Quanto ao método paramétrico, dado o fim analítico e não de previsão do tratamento, não resultou vantajoso utilizar o método Hot Winters ou o Arima.

Algoritmo 1 - Médias Móveis

O método das médias móveis é um método não paramétrico de suavização onde, através de uma operação linear, a série temporal original é transformada na série suavizada³ e onde a ordem da média móvel é o número de unidades temporais (meses no caso) ou movimento médio de N meses a ser considerado no cálculo da função suavizada⁴.

Os dados a serem utilizados neste procedimento podem ser tanto uma série agregada de consumo de cinco ou mais anos quanto os dados completos do cadastro de consumidores (população - 12 meses), dependendo do que se deseja avaliar.

Dados de entrada:

Dados da média de consumo mensal no período analisado, do conjunto de consumidores em questão (ver algoritmo 9 para aplicação no cadastro de 12 meses).

Dados de saída:

Valores transformados pelo Método das Médias Móveis para cada mês, por um número determinado de meses, preferencialmente maior do que 12, de modo que a ordem da Média Móvel varie conforme a análise a ser aplicada.

³ Morettin, P.^a & Toloí, C.M., Métodos Quantitativos: Séries Temporais, 1986, Atual Editora. Brasil.

Procedimento:

Calcula-se uma série de médias, a primeira abrangendo os m primeiros valores da série. A segunda é efetuada da mesma forma que a anterior, mas eliminando a primeira observação e adicionando a observação imediatamente posterior. Continua-se o procedimento até calcular a média das últimas m observações. Cada uma das médias assim obtidas é referente ao instante central do período temporal que limitam¹.

Calcula-se a média móvel dos valores, centrada com a ordem definida através de:

$$X_{kn} = \frac{(X_{n-k} + X_{n-k+1} + \dots + X_n + X_{n+k-1} + X_{n+k})}{k} \quad (01)$$

onde:

X_{kn} = média móvel do n -ésimo elemento

k = é um número inteiro que indica o número de elementos da média móvel, também chamado de ordem da média móvel, que será determinado por $2m+1$ quando for de ordem ímpar e por $2m$ quando for de ordem par.

Dependendo do tipo de flutuação que se deseja eliminar da série de dados original, emprega-se uma ordem diferente para o cálculo da média móvel.

Média Móvel de ordem 2/3 (MM2/3)

A ordem 2/3 é empregada para obter-se a sazonalidade de consumo de uma série de dados de 12 meses.

Dados de entrada:

Dados da média mensal de consumo no período analisado, do conjunto de consumidores em questão (ver algoritmo 9 para aplicação no cadastro de 12 meses) ou dados agregados de consumo mensal.

Dados de saída:

Valores transformados pelo Método das Médias Móveis para cada mês, por 12 meses com média móvel de ordem $k=2$ (onde $m=1$), para ordem par $2m$ e $k=3$ (onde $m=1$), para ordem ímpar $2m+1$.

⁴ Spiegel, M.R., Estadística: Teorya y 875 Problemas Resultados, 1987, McGraw-Hill Editora.

Procedimento:

Aplica-se o Algoritmo 1, tendo, para o primeiro ($n=1$) e para o último ($n=12$) valores $m=1$, ordem par $k=2m$ e para o restante dos dados $m=1$, ordem ímpar $k=2m+1$. Ou seja, para os meses que não o primeiro e o último, calcula-se a média móvel centrada de 3 meses e para os extremos (primeiro e último mês) o valor é obtido a partir da média entre 2 meses apenas (primeiro - segundo e penúltimo - último).

Média Móvel de ordem 3 (MM3)

Para filtragem da instabilidade da data de leitura, foi considerada uma média móvel de ordem ímpar $k=3$ ($m=1, 2m+1$), a qual elimina a variação do comprimento do mês e as irregularidades de leitura. A média móvel de ordem 3 possui a mesma finalidade que a média móvel de ordem 2/3, mas é empregada para uma série temporal mais longa, a série de dados de cinco anos, por exemplo.

Dados de entrada:

Dados da média mensal de consumo do conjunto de consumidores em questão (ver algoritmo 9 para aplicação no cadastro de 12 meses) ou dados agregados de consumo.

Dados de saída:

Valores transformados pelo Método das Médias Móveis para cada mês, por n meses pela média móvel de ordem $k=3$. O vetor resultante é correspondente à série original menos o primeiro e o último elemento.

Procedimento:

Aplica-se o Algoritmo 1 na série de dados, tendo $m=1$, $k=2m+1$ e n igual ao número de dados mensais disponíveis.

Média Móvel de ordem 13 (MM13)

A média móvel de ordem $k=13$ é aplicada para obter-se a componente tendencial e para eliminação da componente sazonal da série temporal original. É empregada em uma série temporal longa, a série de dados.

Dados de entrada:

Dados agregados de consumo mensal do conjunto de consumidores em questão: a série de dados.

Dados de saída:

Valores transformados pelo Método das Médias Móveis para cada mês, por n meses com ordem da média móvel de $k=13$ ($m=6$, $k=2m+1$). O vetor resultante é correspondente à série original diminuído pelos primeiros seis e os últimos seis elementos.

Procedimento:

Aplica-se o Algoritmo 1 na série de dados, tendo $m=6$ e n igual ao número de dados mensais disponíveis.

Algoritmo 2 - Índice de Sazonalidade

O índice de sazonalidade é derivado do cálculo das médias móveis e é utilizado para detectar as variações sazonais de consumo de energia, permitindo que se tenha um valor relativo para viabilizar uma comparação ao longo do período de análise.

O índice de sazonalidade foi calculado pelo método multiplicativo.

Dados de entrada:

Série de dados

Dados de saída:

Índices de sazonalidade percentuais para os meses em que há dados de MM13.

Procedimento:

Aplica-se MM3 e MM13 (Algoritmo 1) para os dados disponíveis, obtendo duas séries de valores mensais denominados por $X_{j,MM3}$ e $X_{j,MM13}$ respectivamente. Calcula-se o índice de sazonalidade para cada mês (IS_j) através da seguinte expressão:

$$IS_j = \frac{X_{j,MM3}}{X_{j,MM13}} - 1 \quad (02)$$

onde:

IS_j = índice de sazonalidade

$X_{j,MM3}$ = valores mensais MM3

$X_{j,MM13}$ = valores mensais MM13

Algoritmo 3 - Método paramétrico

Para o cálculo pelo método paramétrico foi ajustada à série original uma função composta de:

$$\text{Componente de tendência: } f(x) = a + bx \quad (03)$$

$$\text{Componente cíclica } f(x) = c \operatorname{sen}\left(\frac{2p}{12}x + d\frac{p}{2}\right) \quad (04)$$

o modelo assume um ciclo anual, porém, caso haja sazonalidade de 6 meses, como no caso do concessionário de Boa Vista, adiciona-se uma componente senoidal de período $2\pi/6$.

A união das duas funções resulta em:

$$f(x) = a + bx + c \operatorname{sen}\left(\frac{2p}{12}x + d\frac{p}{2}\right) \quad (05)$$

cuja solução, ajustada à série original pelo método dos mínimos quadrados, fornece o valor dos parâmetros da função.

Algoritmo 4 - Ruído Branco

Os problemas na regularidade da leitura criam um distúrbio na série que da teoria de comunicações costuma-se chamar de ruído branco, ruído pois é uma curva aleatória tipo senoidal e branco porque sua distribuição se manifesta em todas as frequências.

A suposição usual é que o ruído branco ou componente aleatória da série temporal seja uma série puramente aleatória, com média zero variância constante². Quando este fato não é observado, esta variação pode ser a fração correspondente a variações que podem estar ligadas a vieses tanto no fornecimento quanto no consumo de energia (variações climáticas inesperadas, características econômicas, calendário de leitura). Considerando que as séries apresentam uma tendência e uma componente cíclica, preferiu-se calcular o ruído branco como quociente entre a série original e a média 3. Com este operador, o ruído branco assume a características de ter média = 1 e variância constante.

Como teste de qualidade da leitura, foi considerado como aceitável uma amplitude da variação de + ou - 5%, correspondente a uma variação de dias de leitura de + ou - 2.

Dados de entrada:

Dados das médias mensais de consumo do conjunto de consumidores em questão (para aplicação ao cadastro de 12 meses, ver algoritmo.8) ou dados agregados de consumo mensal.

O teste do ruído branco pode ser aplicado a toda a população, mas é preferivelmente aplicado separadamente para cada rota, já que existem critérios e rotas de leitura distintas: os grandes consumidores têm em geral, leitura direta por parte do concessionário enquanto os consumidores em baixa tensão são lidos por funcionários terceirizados. Além disso, existe uma diferente qualidade da leitura nas rotas centrais e nas rotas periféricas.

Dados de saída:

Valores do Ruído Branco para cada mês, por um número determinado de meses, preferencialmente maior do que 12, sua média e sua variância.

Procedimento:

Calcula-se a diferença percentual entre a MM3 e os dados reais de consumo para cada mês da série de dados.

$$ruído_{branco} = \frac{X_{ij}}{X_{mn}} \quad (06)$$

onde:

X_{ij} = consumo do mês n

X_{mn} = média móvel de ordem 3, equivalente ao consumo do mês n

Finalmente, calcula-se a média e a variância dos resultados obtidos acima.

Algoritmo 5 - Taxas de crescimento

Índice calculado para verificar crescimento do consumo de energia.

Dados de entrada:

Série de dados de consumo de energia mensal, por categoria e total.

Dados de saída:

Taxas de crescimento do consumo de energia por período predeterminado.

Procedimento:

$$taxa = \frac{\sum_{j=1}^{j=12} X_{Nj} - \sum_{j=1}^{j=12} X_{N-1j}}{\sum_{j=1}^{j=12} X_{N-1j}} \quad (07)$$

onde:

ΣX_{Nj} = somatória do consumo mensal agregado de energia do ano N

ΣX_{N-1j} = somatória do consumo mensal agregado de energia do ano N-1

Algoritmo 6 - Consumo por consumidor

Este dado é obtido das séries agregadas de dados de consumo de energia de um período maior do que 12 meses.

Dados de entrada:

Série de dados agregada de consumo de energia mensal e dados do número de consumidores mensal ou anual.

Dados de saída:

Dados do consumo de energia médio por consumidor.

Procedimento:

Dividir os dados de consumo de energia mensal pelos dados do número de consumidores do mesmo período.

2.2 Cadastro completo dos consumidores (12 meses)

2.2.1 Algoritmos para tratamento primário dos dados

Algoritmo 7 - Filtragem

Filtragem significa partir da população e extrair uma amostra de forma a agrupar os consumidores de acordo com uma característica de consumo comum. Para este estudo, os filtros utilizados para a população serão a classe tarifária e a rota de leitura. Eventualmente, dentro de uma amostra, é possível utilizar-se de uma nova

filtragem para que sejam agrupados consumidores com outras características comuns, como média de consumo ou desvio padrão.

A metodologia de filtragem é extremamente útil no estudo de dados de cadastro de consumidores pois, à medida que se consegue agrupá-los de forma a homogeneizá-los, a análise torna-se cada vez mais precisa.

Dados de entrada:

Dados completos de consumo mensal da população.

Definição dos códigos das classes tarifárias ou definição dos códigos das rotas de leitura.

Dados de saída:

População de consumidores com uma determinada característica comum, por exemplo classe tarifária, rota ou outra filtragem definida.

Procedimento:

"Retira-se" os consumidores da população original segundo cada código de classe tarifária ou rota de leitura.

Algoritmo 8 - Média e desvio padrão do consumo mensal (j)

A média é um valor característico de um conjunto de dados e é classificada como uma medida de posição ou de tendência central. Medida de tendência central por sua vez, é aquela que tende a se localizar em um valor central dentro de um conjunto de dados⁵.

A média dos consumos de energia dos consumidores a cada mês, permite visualizar o perfil de consumo por período ou época do ano.

O desvio padrão é um indicativo da variação que ocorre entre os consumidores, representado na unidade de medida dos valores originais⁶. Este é classificado como uma medida de variabilidade absoluta⁵.

⁵ BONINI & BONINI, Estatística: teoria e exercícios. 1972. LPM Editora

⁶ HAYS, W.L. Statistics for the social sciences. 1973. Holt, Rinehart and Winston Editora. segunda

Dados de entrada:

Séries de consumo mensal do conjunto de dados em questão.

Dados de saída:

Média e desvio padrão do consumo mensal do agregado de dados.

Procedimento:

Calcular a média simples para cada mês (j) através de :

$$\overline{X}_j = \frac{\sum_{i=1}^N X_{ij}}{N} \quad (08)$$

onde:

i = consumidor

j = mês

N= número de consumidores

Calcular a média das médias simples mensais.

$$\overline{X} = \frac{\sum_{j=1}^{12} \overline{X}_j}{12} \quad (09)$$

onde:

j = mês

Procedimento:

Calcula-se o desvio padrão do conjunto de consumidores para cada mês:

$$s_i = \sqrt{\frac{\sum_{j=1}^{12} (X_{ij} - \overline{X}_j)^2}{N}} \quad (10)$$

onde:

i = consumidor

j = mês

N = número de consumidores

Algoritmo 9 - Média e desvio do consumidor (i)

Dados de entrada:

Dados completos de consumo mensal do conjunto de consumidores em questão, por consumidor (X_{ij}).

Dados de saída:

Média anual e desvio padrão dos consumos mensais, por consumidor.

Procedimento:

Calcular a média dos registros mensais por consumidor

$$\overline{X}_i = \frac{\sum_{j=1}^{12} X_{ij}}{12} \quad (11)$$

onde:

i = consumidor

j = mês

Calcular o desvio padrão das médias anuais do consumidor

$$s_i = \sqrt{\frac{\sum_{j=1}^{12} (X_{ij} - \overline{X}_i)^2}{11}} \quad (12)$$

onde:

i = consumidor

j = mês

2.2.2 Desconformidades - teste dos *outliers*

Compatibilidade contratual

As únicas variáveis disponíveis no cadastro da concessionária são o tipo do contrato e a série de consumo histórico portanto, os operadores utilizados realizam verificações nestes dados, oferecendo uma listagem das situações anormais, nas quais é possível que estejam sendo verificadas condições impróprias de alimentação

elétrica. Ocorre notar que esta verificação não necessariamente identifica irregularidade do concessionário, mas em alguns casos pode assinalar irregularidade por parte do consumidor.

Indiretamente os concessionários dispõem de rotinas deste tipo (por exemplo, para validar as leituras dos consumos). Estas não estão disponíveis ao fiscalizador portanto, as aqui propostas servem para integrar aquelas já usadas pelos concessionários.

Algoritmo 10 - Cálculo da Moda

A moda é considerada a medida de tendência central mais simples de interpretar e a mais fácil de calcular. Ela representa o valor ou a classe de máxima frequência dentro dos dados analisados.

Dados de entrada:

Média mensal de consumo no para o período analisado por consumidor do conjunto de dados em questão (algoritmo 9) .

Dados de saída:

Moda das médias mensais do agregado de dados.

Procedimento:

Calcula-se a moda das médias anuais de consumo.

Algoritmo 11- Distribuição de Pareto

O critério adotado neste tipo de algoritmo é o clássico critério dos *outliers*, pelo qual, dada uma população, os membros que se distinguem por qualquer característica, merecem uma especial atenção. Este princípio está relacionado ao princípio de Pareto, este é baseado na premissa de que existe uma distribuição desigual da inconformidade ou viés entre categorias pré definidas e que um reduzido número de categorias em desconformidade com os padrões, é responsável pela maioria dos problemas. A distribuição de Pareto é utilizada geralmente como ferramenta de controle de qualidade em processos produtivos,⁷ própria para tomada de decisão em relação ao estabelecimento de prioridades do tipo: "em que tipo de defeito devemos concentrar nossos esforços?"¹⁵. Esta é uma técnica de estatística descritiva isto é, tem a capacidade de reorganizar os dados sem efetuar nenhuma inferência relacionada

⁷ MONTGOMERY, D.C., Introduction to Statistical Quality Control. 1985. John Wiley & Sons, ⁷ PYZDEK, T. What every engineer should know about quality control. 1989. Marcel Dekker, Inc Inc. Editora. EUA.

aos parâmetros característicos da população em estudo. Para PYZDEK⁸, a definição de Pareto é: *"separar o pouco que é vital do muito que é trivial"*

Propõe-se então, que seja efetuada inspeção aprofundada nestes consumidores. De fato, considerando que o sistema elétrico é composto por milhões de consumidores, a identificação de irregularidades, que com frequência podem representar menos de 0,5% da população, exigiria inspeções muito grandes, difíceis de ser realizadas com profundidade.

Dados de entrada:

Média mensal de consumo no período analisado dos consumidores em questão (aplicação do algoritmo 9 e algoritmo 7), e determinação dos parâmetros de conformidade.

A potência máxima que pode ser demandada nos contratos em baixa tensão é regulamentada pela norma da concessionária, dependendo do tipo de ligação, (monofásico, bifásico e trifásico) e do tipo de transformador que atende ao consumidor (NU001 pp.). Nestes contratos, que atendem uma população homogênea de consumidores, o fator de carga do consumidor pode ser considerado constante, portanto, um aumento do consumo mensal determina um provável aumento da demanda de eletricidade. Desconhecendo a relação efetiva de demanda máxima em função do consumo mensal (fator que varia de região para região), o teste da compatibilidade contratual é realizado ordenando os consumidores (estes já estão agrupados segundo critério de filtragem por classe tarifária) por consumo crescente, calculando a moda, e selecionando os consumidores que se encontram com um consumo superior a 2 ou 3 vezes a moda como consumidores provavelmente mal atendidos (a demanda máxima excede a demanda contratada).

Dados de saída:

Indicação dos valores de consumo em desconformidade (abaixo e acima).

Gráfico de barras e linha dos consumidores crescentemente ordenados pelo parâmetro em questão (médias de consumo, desvio padrão, índice max/média), indicando frequência em valores absolutos da ocorrência de consumidores fora e dentro dos limites estabelecidos, juntamente com apresentação da frequência percentual acumulada da ocorrência dos eventos.

Procedimento:

Definir o limite de conformidade máxima e mínima do parâmetro avaliado, definido para cada caso em estudo (dependente do contrato de fornecimento de energia ou do objetivo da aplicação do algoritmo).

Agrupar em classes os consumidores que apresentam consumo acima, abaixo e dentro do limite de conformidade.

Ordena-se os valores do parâmetro em ordem crescente.

Constrói-se o gráfico destes valores onde se coloca o parâmetro em função dos casos (consumidores), sendo a numeração destes relativas ao total da sub-amostra, ou seja, percentual.

Constrói-se o histograma destes valores.

Elabora-se um relatório com

- valores das faixas do parâmetro analisado,
- contagem de consumidores por faixa e acumulados,
- porcentagem de consumidores por faixa e acumulados.

Algoritmo 12 - Verificação do contrato no setor industrial

Dados de entrada:

Dados completos de consumo mensal da categoria tarifária industrial separada por sub classes.

Dados de saída:

Gráfico contendo patamar de consumo (demanda contratada x 730h) e os consumos registrados no cadastro.

Procedimento:

Calcular energia contratada multiplicando demanda contratada por 730h.

Construir gráfico com linha indicando patamar e os consumos mensais ao longo do ano para cada consumidor.

Regularidade do Faturamento

Para a verificação da regularidade do faturamento, já foi indicada a aplicação do algoritmo 4 - ruído branco, cuja conceituação foi detalhada anteriormente.

Algoritmo 13 - Consumo Mensal Igual a Zero

Dados de entrada:

Dados completos de consumo mensal por consumidor.

Dados de saída:

Indicação do número de meses de consumo igual a zero por consumidor.

Procedimento:

Contar o número de vezes que aparece consumo igual a zero por consumidor.

Algoritmo 14 - Índice Consumo Máximo/Consumo Médio

O índice obtido pela divisão do consumo máximo pelo consumo médio de um consumidor é uma representação da variação de consumo individual.

Dividindo-se o máximo consumo registrado ao longo do período pela média anual de cada consumidor obteremos um índice que, quando for igual a 1 (um), significará que o respectivo consumidor tem seu faturamento registrado por estimativa (os valores serão os mesmos para todos os meses)

Quando este índice apresentar valor superior a 2,5 e não se registrar nenhum consumo mensal igual a zero, estima-se que existe um problema de inconsistência na leitura ou um pico de consumo de um ou dois meses que deve ser analisado. Quando da existência de zeros, o valor alto do índice ocorre justamente por ter sido registrado zero de consumo. O valor que indica conformidade é igual ou menor que 2,5.

Dados de entrada:

Dados completos de consumo mensal do agregado de dados em questão por consumidor.

Dados de saída:

Índice Máximo/Média do consumo mensal por consumidor.

Procedimento:

Aplica-se o Algoritmo 2.

Identifica-se, ao longo do ano, por consumidor, o consumo máximo.

Divide-se os valores de consumo máximo pelos valores de consumo médio calculado.

2.2.3 Distribuição do consumo de energia

Algoritmo 15 - Curva de Lorenz e Índice de Gini

A curva de Lorenz e o índice de Gini são instrumentos utilizados para medir o grau de concentração de uma distribuição, cujo valor varia de zero (perfeita igualdade) até um (a desigualdade máxima).

No caso deste projeto, os dados a serem analisados são aqueles referentes ao consumo médio mensal do setor residencial do cadastro completo de consumidores com registros de 12 meses.

Dados de entrada:

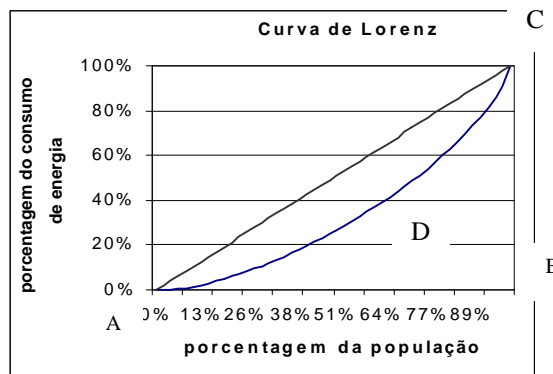
Dados da média mensal de consumo dos consumidores do setor residencial (aplicar algoritmo 7 com filtragem para categoria de consumidores e algoritmo 9).

Dados de saída:

Curva de Lorenz contendo, no eixo x, a porcentagem acumulada das pessoas que consomem até uma determinada quantidade de energia e no eixo y, a porcentagem acumulada do consumo de energia agregado, calculado para cada porcentagem da população.

Índice de Gini

Exemplo



Procedimento:

Ordena-se as médias de consumo em ordem crescente.

Em outra coluna, calcula-se o consumo de energia acumulado e, em seguida, calcula-se a porcentagem de cada valor acumulado em relação ao somatório total.

Cálculo semelhante é efetuado para o número de consumidores.

Exemplo:

	A	B	C	D	E
3	média (consumo por consumidor)	acumulado	%energia	consumidor	%consumidor
4	0	0	0	0	0
5	34	=A5+B4	=B5/\$B\$12	=D4+1	=D5/D12
6	45	=A6+B5	=B6/\$B\$12	=D5+1	=D6/D12
7	46	=A7+B6	=B7/\$B\$12	=D6+1	=D7/D12
8	58	=A8+B7	=B8/\$B\$12	=D7+1	=D8/D12
9	80	=A9+B8	=B9/\$B\$12	=D8+1	=D9/D12
10	140	=A10+B9	=B10/\$B\$12	=D9+1	=D10/D12
11	300	=A11+B10	=B11/\$B\$12	=D10+1	=D11/D12
12	456	=A12+B11	=B12/\$B\$12	=D11+1	=D12/D12

Tem-se a porcentagem acumulada de consumidores e de consumo de energia para construção da curva de Lorenz.

O índice de Gini é igual ao quociente entre a área compreendida entre a diagonal AC e a curva de Lorenz (linha ADC) e a área do triângulo ABC (ver gráfico de exemplo), ou:

$$\text{Índice de Gini} = 1 - \frac{1}{n} \sum_{i=1}^n (\Phi_i + \Phi_{i-1})$$

onde:

n = número de consumidores

Φ_i = proporção acumulada do consumo de energia

$\Phi_i + \Phi_{i-1}$ = somatória da proporção acumulada da energia consumida

Observações

Para a análise estatística dos dados cadastrais, pode-se utilizar qualquer "pacote" estatístico disponível no mercado. Nas análises realizadas neste estudo foram utilizados os seguintes "pacotes": MINITAB[®], STATISTICA[®], EXCEL[®].

Caracterização dos Dados para Aplicação dos Algoritmos



Convênio ANEEL/FUSP: Estudos e Pesquisas Sobre Regulação e
Fiscalização do Setor Elétrico Brasileiro.
Projeto: Metodologia de Análise de Cadastros de Concessionárias

Índice

Introdução	3
1. Relatório a ser solicitado ao concessionário	3



Convênio ANEEL/FUSP: Estudos e Pesquisas Sobre Regulação e
Fiscalização do Setor Elétrico Brasileiro.
Projeto: Metodologia de Análise de Cadastros de Concessionárias

Introdução

Este volume 2 do manual III contém as indicações sobre o tipo de dados que a ANEEL poderá solicitar aos concessionários para elaboração dos tratamentos especificados.

1. Relatório a ser solicitado ao concessionário

Em virtude da facilidade de introdução de rotinas matemáticas às rotinas usualmente efetuadas pelo concessionário, tanto para uso interno como para finalidade fiscalizadora, existe a opção para o órgão fiscalizador de solicitar apenas os relatórios com o resultado da aplicação dos algoritmos.

No caso do órgão fiscalizador optar por efetuar as análises, recomenda-se que este solicite os seguintes dados:

1. Série longa de dados de consumo, no mínimo 5 anos, agregados por classe tarifária.
2. Cadastro Comercial completo do Concessionário:
 - série de consumo de 12 meses, onde os dados devem conter as seguintes informações de cada consumidor:
 - a. Identificação
 - b. Local
 - c. Setor
 - d. Rota
 - e. Classe
 - f. Subclasse
 - g. Tensão
 - h. Ligação
 - i. Bairro
 - j. Número de Medidores -Consumidor
 - k. Leitura do consumidor nos últimos 12 meses
 - l. Código de atividade
 - m. Demanda Contratada
 - n. Demanda Prevista
 - o. Fator de Carga Típico

A seguir estão indicadas o tipo de dados necessários para a aplicação de cada algoritmo proposto no relatório II.

GRUPO I

<i>algoritmo</i>	<i>tipo de dado necessário para aplicação</i>
Algoritmo 1 - Médias Móveis	série de dados do consumo de 5 anos ou mais
Algoritmo 2 - Índice de Sazonalidade	série de dados do consumo de 5 anos ou mais
Algoritmo 3 - Método paramétrico	série de dados do consumo de 5 anos ou mais
Algoritmo 4 - Ruído Branco	série de dados do consumo de 5 anos ou mais
Algoritmo 5 - Taxas de Crescimento	série de dados do consumo de 5 anos ou mais
Algoritmo 6 - Consumo por Consumidor	série de dados do consumo de 5 anos ou mais

Os algoritmos de cálculo de sazonalidade e tendência demandam uma série longa de dados, no mínimo 5 anos, mas com relação ao número de consumidores avaliados, recomenda-se a filtragem em grupos homogêneos, por exemplo, por macrorregião geográfica abrangida pela concessionária e/ou por setor de consumo, eventualmente subdividido por classe tarifária.

GRUPO II

<i>algoritmo</i>	<i>tipo de dado necessário para aplicação</i>
Algoritmo 7 - Filtragem	aplicável tanto para séries longas e dados agregados como para o cadastro completo de consumidores
Algoritmo 8 - Média e desvio padrão do consumo mensal (j)	cadastro completo de consumidores
Algoritmo 9 - Média e desvio do consumidor	cadastro completo de consumidores

GRUPO III

<i>algoritmo</i>	<i>tipo de dado necessário para aplicação</i>
Algoritmo 10 - Moda	cadastro completo de consumidores
Algoritmo 11 - Distribuição de Pareto	cadastro completo de consumidores
Algoritmo 12 - Verificação do Contrato no setor Industrial	cadastro completo de consumidores
Algoritmo 13 - Consumo Mensal Igual a Zero	cadastro completo de consumidores
Algoritmo 14 - Índice Consumo Máximo/Consumo Médio	cadastro completo de consumidores
Algoritmo 15 - Curva de Lorenz e Índice de Gini	cadastro completo de consumidores

Os algoritmos para verificação da adequação contratual e regularidade do consumo devem necessariamente ser aplicados em populações homogêneas. No caso da adequação contratual, recomenda-se que o concessionário envie relatório indicando a distribuição do consumo, por faixa de consumo com média anual ou dos 3 meses mais frios e mais quentes, de todos os consumidores agrupados por classe tarifária. Para o cálculo do ruído branco, o concessionário deverá aplicar o algoritmo para todas as rotas de leitura e o relatório deve evidenciar o número de rotas que apresentou índice maior que 10% de variação.

Junto com a aplicação do algoritmo para verificação da adequação contratual, o concessionário deve aplicar para as populações de consumidores das diferentes classes tarifárias, a rotina matemática de cálculo do índice máx/média. O relatório deve indicar:

- o número de consumidores com índice máx/média $> 2,5$ e sem a presença de zero.
- o número de consumidores com índice máx/média = 1.
- o número total de consumidores.



Séries Temporais



Convênio ANEEL/FUSP: Estudos e Pesquisas Sobre Regulação e
Fiscalização do Setor Elétrico Brasileiro.
Projeto: Metodologia de Análise de Cadastros de Concessionárias

Índice

Introdução	3
1. Classificação das séries temporais	3
2. Objetivos das análises de séries temporais	5
3. Métodos de Análise	7
4. Exemplos de séries temporais na indústria elétrica	7



Convênio ANEEL/FUSP: Estudos e Pesquisas Sobre Regulação e
Fiscalização do Setor Elétrico Brasileiro.
Projeto: Metodologia de Análise de Cadastros de Concessionárias

Introdução

Este apêndice do manual 3 contém texto sobre a conceituação de séries temporais, baseado na publicação *Manual para la Realización de Estudios sobre Usos Finales de Electricidad, Estudio del Mercado*, elaborado por Alessandro Barghini, São Paulo, 1996.

Uma série temporal é definida como um conjunto de dados ordenados em relação à variável principal representada pelo tempo. São séries temporais, por exemplo, os valores registrados diariamente de uma determinada ação na bolsa de valores ou os valores de temperatura registrados em uma determinada localidade ao longo do tempo. Citando exemplos elétricos, são séries temporais os valores de corrente registrados com um amperímetro em uma subestação, ou os registros dos consumos de energia faturados, mês a mês, de um consumidor ou de um conjunto de consumidores. De certa maneira, é possível afirmar que grande parte dos dados de que dispõe uma empresa elétrica está organizada na forma de séries temporais.

O estudo das séries temporais é portanto, um instrumento extremamente poderoso para o conhecimento do mercado elétrico e qualquer operador que se dedica ao planejamento elétrico pode obter grandes benefícios do conhecimento das técnicas de análise das séries temporais.

1. Classificação das séries temporais

De um ponto de vista estritamente matemático, podemos usar diferentes critérios de classificação das séries temporais. A primeira classificação é em relação à continuidade do fenômeno e pode ser separada em dois grandes grupos: séries contínuas e séries discretas.

Séries contínuas: são conjuntos de dados referentes a eventos que se apresentam como uma função contínua. Uma série contínua é por exemplo, o gráfico de um amperímetro de pena que registra os valores instantâneos da corrente de uma determinada carga. A maioria das séries temporais são, por sua própria natureza, séries contínuas, já que se realizam permanentemente no tempo. O fato de uma série ser contínua, não significa que ela tenha que ser tratada como tal e, de fato, a maioria das séries contínuas tende a ser representada de forma descontínua já que as observações são espaçadas no tempo. Por exemplo: os registros de corrente das subestações são marcados de hora em hora, os valores de consumo nos medidores

de faturamento hora sazonal são, em geral, integrados de 15 em 15 minutos, os registros de consumo dos consumidores são geralmente armazenados mês a mês.

Se consideramos Δt um intervalo constante de registro de um fenômeno contínuo e, a cada intervalo realizamos uma leitura, transformamos automaticamente uma série contínua em uma série discreta, a qual terá $N = T/\Delta t$ pontos, onde T é o tempo total durante o qual se verifica ou é registrado o evento e Δt é o intervalo de medição.

Séries discretas: são conjuntos de dados ordenados, selecionados em intervalos discretos que podem ser constantes ou variáveis.

Aqui vai um alerta sobre o processo de transformação de um fenômeno contínuo em uma série temporal descontínua: como em todos os processos de medições precisas, precisamos sempre verificar a leitura, entender os desvios introduzidos pela leitura e controlar as unidades de medida utilizadas. Um dos problemas criados pela leitura é o erro hoje em dia cometido com as medições realizadas com medidores eletrônicos acumuladores de consumo: é freqüente encontrar a integral do consumo em um intervalo de 15 minutos como potencial do sistema durante o intervalo sendo que, na realidade, trata-se de uma integral de consumo que, dividida pelo tempo, oferece uma potência média e, em muitos estudos específicos, especialmente nos estudos dos transformadores, é um valor inadequado para um estudo de dimensionamento.

A segunda classificação das séries temporais é em relação a quantidade de variáveis medidas. Neste ponto de vista, elas podem ser divididas em 4 grupos básicos: séries monovariadas, séries multivariadas, séries multidimensionais e séries multivariadas multidimensionais.

Serie monovariada: dada uma série, ela será monovariada quando, ao longo da variável independente tempo existe o registro de uma única variável, por exemplo o consumo de um consumidor, ou o registro da potência ou do consumo de um transformador. Portanto, a série assumirá a forma de:

$$f(t)x = x_0 \ x_1 \ \dots \ x_{n-1} \ \dots \ x_n$$

na qual x_0 representa o valor da variável dependente x no instante t_0 e assim por diante.

Série multivariada: dada uma série temporal, ela é multivariada quando, ao longo da variável independente tempo, existem os registros de mais variáveis do mesmo tipo. Por exemplo, uma curva de carga poderia ser considerada como a

somatória das curvas de demanda de cada um dos consumidores, mais a curva de perdas. Uma curva total, por exemplo de uma subestação, pode ser considerada como a somatória das curvas seus alimentadores, cada um dos quais possuindo um perfil bem característico de consumo, sendo o primeiro com consumo predominantemente industrial, o segundo com consumo predominantemente comercial e o terceiro com consumo predominantemente comercial. Podemos ter, então, uma curva global que é representada pela somatória das três séries, a série $f(x)_{icr}$ assumirá o desenvolvimento do tipo

$$f(t)_i = X_{i0} \quad X_{i1} \quad \dots \quad X_{i(n-1)} \quad X_{i(n)}$$

$$f(t)_c = X_{c0} \quad X_{c1} \quad \dots \quad X_{c(n-1)} \quad X_{c(n)}$$

$$f(t)_r = X_{r0} \quad X_{r1} \quad \dots \quad X_{r(n-1)} \quad X_{r(n)}$$

Onde a primeira série $f(t)_i$ representa a série de valores de demanda do alimentador industrial, a série $f(t)_c$, a série de valores de demanda do alimentador comercial, a série $f(t)_r$, a série de demanda do alimentador residencial.

Série multidimensional. Se, à variável independente tempo correspondem mais valores com dimensões diferentes, a série é definida como uma série multidimensional, já que, dependendo da série, ela assume uma dimensão diferente. No setor elétrico, séries típicas multidimensionais são as séries de registros de corrente de um consumidor, de um transformador, de uma subestação ou até do sistema, e outra variável, como por exemplo a temperatura ambiente, o nível de intensidade luminosa ou o nível de atividade econômica.

Série multidimensional multivariada. Se, à variável independente tempo correspondem mais valores, alguns com a mesma dimensão/unidade e outros com dimensões diferentes, a série é definida como uma série multidimensional e multivariada já que existe mais de uma série com a mesma dimensão e existem uma ou mais séries com dimensões diferentes. No setor elétrico, séries típicas multidimensionais e multivariadas são as séries de registros de corrente de diferentes consumidores, como o exemplo das 3 séries de 3 alimentadores diferentes e de uma ou mais outras variáveis dependentes, como por exemplo a temperatura ambiente, o nível de intensidade luminosa ou o nível de atividade econômica.

2. Objetivos das análises de séries temporais

A finalidade última da análise das séries temporais é a de ordenar uma quantidade ampla de dados e, a partir desta ordenação, alcançar uma série de

objetivos os mais variados possível. A grande variabilidade dos objetivos é resumida por Turkey em 3 objetivos gerais:

1. Criar um modelo do fenômeno observado que permita descrever a variabilidade. Este primeiro objetivo pode também ser descrito como a criação de um modelo que permita controlar o processo. Neste sentido, este aspecto é especialmente importante na indústria elétrica para a elaboração de modelos de despacho de cargas e, em geral, para dimensionamento das instalações (disponibilidade de potência).

2. Extrair conclusões estatísticas sobre o fenômeno, principalmente em termos de repetição das componentes observadas. Este objetivo é relevante para a empresa elétrica, especialmente no planejamento de manobras de máquinas e de despacho de cargas.

3. Realizar previsões, principalmente no que diz respeito a possíveis trajetórias futuras. Este objetivo é especialmente importante para a empresa elétrica no planejamento do sistema.

Dentro dos objetivos gerais, é possível destacar uma série de objetivos específicos de análise das séries temporais na indústria elétrica:

- obter os valores de carga ou de demanda máxima dos transformadores de distribuição em função dos consumos mensais de agregados de consumidores.
- Obter os valores de demanda por consumidor e por conjunto de consumidores, por eletrodoméstico.
- Prever a repetição do comportamento da demanda a nível de subestação ou de sistema ao modificar variáveis macroeconômicas ou ambientais.
- Prever o comportamento do consumo em função do nível de atividade econômica.
- Realizar projeções de consumo e de demanda em períodos futuros, a partir de variáveis macroeconômicas.

A lista de objetivos poderia se alongar muito e, de fato, pode-se afirmar que poucos métodos de cálculo têm uma aplicação tão ampla na indústria elétrica como a análise das séries temporais. Como pode ser visto pela lista, o interesse pelos resultados das análises das séries temporais cobre praticamente todos os setores da indústria, como: distribuição, tarifa, despacho de cargas e planejamento. Praticamente todo o universo de atividade da indústria elétrica pode ser influenciado pelos resultados das análises das séries temporais.

3. Métodos de Análise

Dentro da grande variabilidade de objetivos de análise das séries, é possível utilizar diferentes métodos que podemos sumariamente classificar em 3 grupos:

Unitária: Considera a série inteira como uma das possíveis realizações de um processo estocástico⁹ que pode realizar-se no tempo. Neste caso, o interesse na análise da série é verificar a condição de repetição da seqüência, isto é, encontrar um padrão de comportamento dos valores como o que ocorre quando alisamos a curva de carga de uma série de dias, procurando montar as condições de repetição da curva nos diferentes dias da semana.

Desagregada: Considera cada ponto da série como a realização possível do evento dentro de uma determinada distribuição de probabilidade da variável aleatória. Isto ocorre por exemplo, quando analisamos a demanda máxima em uma determinada hora e correlacionamos este valor com variáveis como temperatura, índice de luminosidade, umidade relativa ou outras variáveis atmosféricas procurando identificar de que forma uma destas variáveis aleatórias afeta a demanda elétrica. Neste caso, ao lado de uma série principal função de T, a demanda elétrica, registraremos uma segunda série de dados representada pela segunda variável, a temperatura. A série será portanto, multivariada.

Agregada: Finalmente, pode-se também considerar uma série como a somatória de séries distintas e estudar o comportamento da variável determinada pela influência de seus componentes.

4. Exemplos de séries temporais na indústria elétrica

Dentro da indústria elétrica encontramos um número grande de exemplos de séries temporais. Apesar de todas se referirem ao comportamento do fenômeno elétrico ao longo do tempo, elas representam características distintas e merecem tratamentos distintos. Entre os exemplos clássicos de séries temporais encontramos:

- registros gráficos de corrente de consumidores individuais ou de conjunto de consumidores, por exemplo de transformadores.
- registros gráficos de corrente ou de energia de alimentadores.
- registros horários de subestações ou do sistema.

⁹ processo estocástico é aquele que não permite um cálculo exato dos eventos, mas sim a probabilidade de cálculo deste dentro de limites especificados (BOX & JENKINS, *Time serie analysis: forecasting and control*. 1976.

- registros de faturamento dos clientes, que podem ser individuais, consumidor por consumidor ou agregados, quando a série soma o comportamento de grupos homogêneos de consumidores, por exemplo por faixa tarifária ou por categoria de consumo.

Junto destas séries *elétricas* podemos dispor de outras séries de diferentes naturezas. As séries temporais típicas utilizadas pela indústria elétrica são:

- séries econômicas, como nível de atividade econômica, nível de renda.
- séries meteorológicas, como registros de temperatura, de duração do dia, velocidade do vento e similares.

Estas séries, como pode ser imediatamente observado, possuem características distintas. A primeira, dos registros de corrente dos consumidores, é tipicamente uma série contínua. A segunda, dependendo do tipo de instrumento, pode ser uma série contínua (se a medição for gráfica) ou uma série descontínua se é usado um medidor eletrônico que armazena os valores de energia ou de potência, geralmente de 15 em 15 minutos. A diferença principal entre as duas séries é que a primeira é uma curva não diversificada, onde é possível ler individualmente os eventos elétricos que determinam uma determinada demanda e, com um estudo atento, podemos em teoria reconstruir as componentes que determinaram a demanda. A segunda é uma curva composta pela somatória do comportamento de diferentes consumidores e seu tratamento deverá ser estatístico, já que o comportamento individual se dilui no ruído de fundo do sistema (a influência do consumidor individual é inferior à sensibilidade do instrumento). A terceira, a série dos registros das subestações, costuma ser uma série descontínua, em geral com intervalo de um ou mais meses.

Equipe do projeto:

Prof. Ildo Luis Sauer

Mariana de Oliveira Pedreira

Paulo Marcos Santo de Almeida

Alessandro Barghini

André Xavier Lima